REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)* 30-09-2007	**2. REPORT TYPE** Final Report	**3. DATES COVERED** *(From - To)* 01-09-2006 to 30-09-2007
4. TITLE AND SUBTITLE Geometric Methods for ATR: Shape Spaces, Metrics, Object/Image Relations, and Shapelets		**5a. CONTRACT NUMBER**
		5b. GRANT NUMBER FA9550-04-1-0302
		5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S) Dr. Peter F. Stiller Institute for Scientific Computation, MS 3404 Blocker Bldg. 612 Texas A&M University College Station, Texas 77843-3404		**5d. PROJECT NUMBER**
		5e. TASK NUMBER
		5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Texas Engineering Experiment Station 332 Wisenbaker Engineering Research Center Texas A&M University College Station, Texas 77843-3000		**8. PERFORMING ORGANIZATION REPORT NUMBER**
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) Air Force Office of Scientific Research 4015 Wilson Boulevard Arlington, Virginia 22203-1945 Dr Sjogren/NE		**10. SPONSOR/MONITOR'S ACRONYM(S)** AFOSR
		11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT

Distribution Statement A: unlimited

AFRL-SR-AR-TR-07-0465

13. SUPPLEMENTARY NOTES

14. ABSTRACT

This report details our recent progress (1 Sept. 2006 to 30 Sept. 2007) and summarizes the work done over the entire 40 months of the project. It outlines our recent success in creating the first "global" object/image equations and metrics for the full perspective case, and it discusses some of the additional experimental testing we did to verify the robustness of our algorithms. We extended our analysis of the relationship between the shape spaces for point features under similarity transformations and those under the affine group. The former are relevant to orthographic sensors (radar) and the later arise when dealing with weak perspective sensors (optical – far field). Understanding the relationship between the two types of shape spaces facilitates fusing data from these two types of sensors. In addition, we worked on global embeddings of the shape spaces in the orthographic (radar) case, and we continued to work on metrics for the case where the features are taken as unordered collections of points. Currently, we are working on the 3D reconstruction problem, especially shape from motion and have started to look at advanced statistical/noise issues.

15. SUBJECT TERMS

target recognition, object/image equations, object/image metrics, shape spaces, shape statistics, invariants, shapelets, 3D shape reconstruction

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON** Dr. Peter F. Stiller
a. REPORT U	**b. ABSTRACT** U	**c. THIS PAGE** U	UU	64	**19b. TELEPHONE NUMBER** *(include area code)* 979-862-2905

Final Report
Air Force Office of Scientific Research
Grant No. FA9550-04-1-0302

"Geometric Methods for ATR: Shape Spaces, Metrics, Object/Image Relations, and Shapelets"

Dr. Peter F. Stiller, Principal Investigator
Professor of Mathematics and Computer Science
Institute for Scientific Computation
Texas A&M University
College Station, Texas 77843-3404
stiller@math.tamu.edu or stiller@isc.tamu.edu

Ms. Christian D. Stratta, Senior Project Administrator
Office of Sponsored Research
Texas Engineering Experiment Station
332 Wisenbaker Engineering Research Center
Texas A&M University
College Station, Texas 77843-3000
c-stratta@tamu.edu (979) 458-7623

September 30, 2007

Objectives: This three-year research effort is being conducted at Texas A&M University by the principal investigator, Dr. Peter F. Stiller, and a number of graduate research assistants.

We begin this report by reviewing the project's objectives as outlined in the original proposal abstract.

The general problem of single-view recognition is central to many target recognition and computer vision tasks. For example, efficiently recognizing three dimensional arrangements of features (such as geometric configurations of lines and/or points) from a single two dimensional view is a key research question. A solution will require an approach that is view and pose independent. Unfortunately, existing methods often rely on computationally expensive template matching that is, strictly speaking, not view or pose invariant. Instead those methods use comparisons against templates created for each possible view; with the infinite number of possibilities being approximated by some finite number of views. To carry out an invariant approach to target recognition, we must seek out properties and relationships that are geometrically intrinsic to the objects and/or images being compared. We must also be aware that different sensors necessitate different models of image formation and therefore different forms of invariance. Radar and Ladar make use of an orthographic model, while most optical sensors use a weak perspective or full perspective model.

Consideration of view and pose independence, as well as a desire for a coordinate independent formulation, leads naturally to characterizing configurations of features by their 3D or 2D geometric invariants. The specific group (Euclidean group, similarity or conformal group, affine group, or projective general linear group) to which things should be invariant is a function of the sensor type. We also need to determine a fundamental set of equations that express the relationship between the 3D geometry and its "residual" in a 2D (or 1D) image. These are known as object/image equations. They should completely and invariantly describe the mutual 3D/2D constraints. Once found, they can be exploited in a number of ways. For example, from a given 2D configuration, one could use the OI-equations to derive a set of nonlinear constraints on the geometric invariants of the 3D configurations capable of producing that given 2D configuration, and thus arrive at a test for determining the object being viewed. Conversely, given a 3D geometric configuration (features on an object), one could derive a set of equations that constrain the invariants of the images of that object; helping us determine if that particular object appears in various images.

We propose to create so-called "global" forms for the object/image equations, study their properties (especially under geometric degeneracy), and exploit them to develop new algorithms for target recognition. This will require using advanced mathematical techniques from algebraic and differential geometry to construct generalized shape spaces for various projection and sensor models. We will use that construction to find natural metrics that express the distance (difference) between two object configurations, two image configurations, or an object and an image pair.

These metrics should produce the most robust tests for target identification; at least as far as target geometry is concerned. Moreover, such metrics will provide the basis for efficient hashing schemes to do target identification quickly, and they will provide a rigorous foundation for error analysis in the ATR process.

A summary list of the research topics that are included in this proposal appears below:

Proposed Tasks and Problems

1. Object/Image Relations and Shape Spaces
2. Extending the O/I Formulation to Other Feature Sets and Sensor Models
3. Symbolic Computation and Alternative Methods for Computing the O/I Equations
4. Metrics
5. Recognizing Articulated Objects
6. Geometric Hashing
7. Unlabeled Matching
8. Shapelets
9. Noise Analysis
10. Performance Prediction
11. Technology Transfer

Status of Effort: (Period of Performance 6/1/04 to 9/30/07.)

At the time of this writing the grant has ended, having been on-going for the previous 40 months.

Recall that in previous AFOSR sponsored work we were able to achieve several important results, including the understanding, development, and analysis of a global approach to invariants and object/image equations in the generalized weak perspective (affine) case. That work also included our initial construction of a new class of discrimination metrics that are generalizations of the classical Procrustes metric of statistical shape theory. In the first instance, we provided a complete dictionary between the old algebraic approach to invariants and the new, more geometric, global approach. This was worked out completely in the generalized weak perspective case and appears in our paper "Object/Image Relations, Shape Spaces, and Metrics" and more recently in a book chapter entitled "Object-Image Metrics for Generalized Weak Perspective Projection," in *Statistics and Analysis of Shapes,* edited by Hamid Krim and Anthony Yezzi, Jr. and published in 2006 by Birkhauser. This new approach creates a geometric framework for discrimination theory and a more robust approach to recognition. Some of the main ideas and their application to the full perspective (optical) case were presented in our paper, "Global Invariant Methods for Object Recognition" described in a previous report. New results on this topic have just appeared in our paper "Recognizing point configurations in full perspective," which was written jointly with our graduate student Kevin Abbott for the Electronic Imaging Conference, Vision Geometry XV.

Overall our global approach provides a way to explore the behavior of recognition algorithms when dealing with multiplicities or geometric degeneracies (which cannot be handled with other methods). The difficulty in using the classical numerical invariants for this purpose is that they are only rational functions on the appropriate quotient variety. As such, they are not always defined. This leads to serious numerical problems in any algorithm based on these invariants. To remedy these problems, we succeeded in replacing these invariants by points in a Grassmann manifold in the weak perspective case, or by certain geometric objects, namely toric sub-varieties of Grassmannians, in the full perspective case. The object/image equations become the expression of certain incidence relations in the weak perspective case or, in the full perspective case, certain "resultant-like" expressions for the existence of a non-trivial intersection of the toric sub-varieties with certain Schubert varieties in the Grassmannian. This "global" approach to invariants is providing more robust object recognition algorithms. Moreover, by representing the relevant shape spaces as varieties embedded in projective space, we can endow each shape space with a metric by restricting the standard Fubini-Study metric. These ideas are discussed in our paper "Object Recognition from a Global Geometric Perspective - Invariants and Metrics." This approach produces a natural metric on both the object and the image space that can be exploited to create an effective discrimination theory (i.e. a meaningful notion of "distance" between objects, between images, and between an object and an image.) Finally, several new directions have emerged from our work. These directions have been incorporated into our research and include the study of object/image equations for unordered point features to facilitate point cloud matching, research on object/image equations with parameters to handle articulation of objects, the investigation of invariant point to surface matching, 3D reconstruction from motion, and the statistics of shape for noise analysis. Progress on these will expand the recognition power of our approach and its applicability to Air Force problems.

Accomplishments/Findings:

We report below (in summary form) on several significant areas of progress. Details can be found in the listed papers.

1. Shape Statistics

Kendall pioneered statistical shape theory for point features in the plane under similarity transformations. Among his results is a description of the distribution of shapes for point features selected from independent spherical normal distributions each with covariance matrix normalized to the 2 by 2 identity matrix and with means at selected points in the plane. One can regard this as an early attempt to introduce the idea of "noisy" shapes. An important question is to determine for a given distribution of object shapes, the corresponding distribution of image shapes under appropriate hypotheses. This was something not addressed by Kendall or others working in this area. Building on Kendall's results in 2D, we are trying to answer the above question in a particular case involving a small number of point features in the plane under similarity transformations which are projected to 1D. This is a modified radar case where scale is unknown.

2. 3D shape reconstruction from motion.

This is the newest aspect of our work. The goal is to improve upon the ideas and methods of Mark Stuff to do 3D target geometry reconstruction from a series of 1D radar range profiles taken of a target moving relative to the sensor. We have been working with a simplified 2D to 1D orthographic model which captures the essence of the problem. The central issue is how to find the object that best fits the data provided by the accumulated set of noisy images. In our formulation this means finding the object whose image locus in the image shape space best fits the image data. The key is what is meant by "best fit." We argue that the natural Riemannian metric on the image shape space is the best measure to use in determining "goodness of fit," and we are attempting to design an optimal fitting procedure based on this idea.

3. Global descriptions of shape spaces in the orthographic (radar) case.

In order to carry out our program of developing the global version of the object/image equations and object/image metrics for the orthographic case it is necessary to understand how the shape spaces for points features in this case isometrically embed in standard Euclidean space. For small numbers of point features in 1D this is relatively easy, but for greater numbers of points in 1D and any number in 2D or 3D, this becomes a harder problem. It essentially amounts to finding an embedding of real or complex projective space isometrically into a Euclidean space (real or complex) of as low a dimensional as possible, and then extending this embedding to a certain cone over the projective space. We have been able to do this, paving the way for the full development of our approach to recognition in the radar case.

4. Testing our algorithms and new applications.

Work on designing and implementing experiments to test several recognition algorithms based on our object/image metrics was carried out during the course of the project. The results were summarized in our paper, "Robustness and statistical analysis of object/image metrics," which was presented at Electronic Imaging, Vision Geometry XIV, in San Jose, CA, in January 2006. The results showed that for point features in the weak perspective case, our object/image metric performs surprisingly well even in the face of sensor noise. The results also scaled well with respect to the size of the object database and showed the expected strong increase in target matching performance with each additional feature point considered. In addition, we have recently been collaborating with Ms. Olga Mendoza, a young researcher at AFRL, Wright –Patterson AFB, who has performed additional tests of the algorithms, and who has an interest in applying the object/image metrics to problems in image registration and tracking.

5. The Full Perspective Case – O/I Equations and Metrics

In joint work with our Ph.D. student Kevin Abbott we have made significant progress in the very difficult case of full perspective projection (essentially the pin-hole camera model of projection) which is important for recognizing objects in optical images. The central difficulties in this case are that the shape spaces are not well understood and that the computational complexity increases dramatically when dealing with projective invariance. We were able to make significant progress this past year, introducing the first true global version of the object/image equations in the full perspective case, and the first metrics fully invariant to projective/perspective transformations. These results appear in our paper "Recognizing Point Configurations in Full Perspective," and in greater detail in Kevin Abbott's Ph.D. thesis.

6. Recognizing configurations of linear features in the generalized weak perspective case.

Very little previous work has been done on the shape theory of line configurations. In the course of this project we carried out an investigation of the problem of single-view recognition for sets of line features under generalized weak perspective projection. In particular we derived the object/image equations for projection from 2D to 3D. Unfortunately because this required using the Plucker coordinates of the lines, we had to fall back on standard position methods to define our invariants, meaning that the results require certain general position assumptions that we would eventually like to eliminate. In addition, this work on the generalized weak perspective case has revealed an approach to the orthographic (radar) case which we hope to flesh out shortly. Our results appear in a paper entitled "Recognition of Configurations of Lines I – Weak Perspective Case" which was published in late 2005.

7. Comparison between shape in the similarity case and shape in the affine case.

The definition and study of shape spaces for the similarity group began with Kendall in 1977. He treated ordered k-tuples of points in Euclidean m-space (not all the same point). Two such k-tuples of feature points (or "landmarks") are deemed equivalent if they differ by a similarity transformation (i.e. rotation, translation, and/or positive scale). The resulting equivalence class is known as the "shape" of the configuration of the k feature points. The geometry of these spaces has been studied by many authors. In previous work we examined the shape spaces for the larger affine group and explored the relationship between the shape of a configuration of points in three dimensions and the shapes of all the images of that configuration in two dimensions under all possible (affine) generalized weak perspective projections. This leads to the notion of the object/image equations which quantify the relationship between 3D object features (points) and 2D image features. They are zero if and only if a generalized weak perspective projection exists which takes the 3D data to the 2D data. The geometry in this case is particularly nice, relating as it does to properties of Grassmann manifolds. Also the natural metric geometry, both in the classical similarity case (Procrustes metric) and the affine case (Fubini--Study metric), provides a way to measure the distance between shapes, both object shapes and image shapes, as well as providing a natural notion of distance (i.e. matching) between an object shape and an image shape.

In our initial work under this grant, we sought to gain a clearer understanding of the various notions of shape (i.e. shape for different transformation groups acting on the feature points). The first problem we considered was the relationship between the shape spaces of Kendall for the action of similarity transformations, consisting of rotations, translations and scale (no reflections for the moment) and the shape spaces for the action of the affine group. We were able to make a complete analysis of this situation. The primary discovery was the rather simple and elegant form of the map taking you from similarity shape to affine shape and the complete analysis of the locus of degenerate similarity shapes which leads to some interesting topological issues.

The results of this work appear in our paper "The Relationship Between Shape under Similarity Transformations and Shape under Affine Transformations," a copy of which was attached to one of our previous reports.

8. Invariants and Shape Characterizations for Unordered Feature Points

We have begun investigating extensions of our methods to the difficult case of unordered feature points. Here we can make use of new work of Boutin and Kemper, "On Reconstructing Configurations of Points in the Projective Plane from a Joint Distribution of Invariants," preprint, April 2004. This paper provides a complete description of the invariants for unordered point features in an image under full perspective. Our hope is to extend this to object features in 3D and then combine it with the projection from 3D to 2D to obtain invariant matching equations (object/image equations) that are also permutation invariant. This would in turn hopefully lead to permutation invariant metrics for point feature object/image comparisons in the full perspective case.

9. The Weak Perspective Case for Point Features

This material has been reported on previously and was recently written up in an invited book chapter (see below) published by Birkhauser.

Characteristics of our Results

Below is a brief outline in bullet form of the principle characteristics of our approach to using object/image metrics for sensor exploitation and target identification.

+ Is a Feature Based Approach to Target Identification

The approach makes use of small numbers of sensed features associated with features in the target geometry.

+ Invariance to Pose and Scale

The method allows identification to be achieved across all poses and, if desired, at varying scales without resorting to exhaustive template matching.

+ Based on Intrinsic Measures of Shape

To achieve invariance the method makes use of the emerging mathematical theory of shape that characterizes internal relationships among features, independent of relevant transformations like rotation, translation, and scale. These characterizations turn out to also be independent of the coordinate system used to record the target or image feature locations.

+ Permits an Invariant Characterizations of the Fundamental Matching Criteria known as the Object/Image Equations

We can express the necessary and sufficient conditions for the invariant shape of a set of target features to be consistent with the invariant shape of a set of image features as a set of equations in local or global coordinates on the space of object shapes and image shapes. Consistency in this case means that there exists some pose of the object, some sensor location, and some set of sensor parameters that will result in the target features projecting through the sensor to the observed image features. These equations are called object/image equations.

Since the input to the (non-linear) object/image equations are a set of target shape coordinates and a set of image shape coordinates, we can use the equations to invariantly determine matching and also to determine all image shapes that can be achieved from a given object shape or all object shapes capable of producing a given image shape.

+ Relies on Metric Geometry

The methods yield intrinsic metrics on the "spaces of shapes". These metrics satisfy the usual triangle inequality and can be used to measure object or image shape similarity (up to the allowable transformations, e.g. rotation, translation, scale, etc.) In addition they provide a mechanism for effective hashing in large databases of target or image feature sets.

+Yields a Natural Measure of Matching (Distance) between an Object Feature Set and an Image Feature Set

The theory provides two natural "metrics" for invariantly matching a given set of object features to a given set of image features. We can compute in object space, using the metric distance between object shapes, by finding the minimum distance between the given object and all objects capable of producing the given image. The alternative is to work in image space, using the metric in that space, to compute the minimum distance between the given image and all images of the given object. A deep duality theorem assures that, with suitable normalization, these two metrics are the same. This means that for any particular sensor type amenable to this approach, there is a unique natural pose and scale invariant measure of object/image closeness of match!

+Amenable to Statistical Analysis and ATR Theory

Shape was originally introduced for the purpose of doing invariant statistical analysis. Recent workshops at AIM (American Institute for Mathematics), IMA (Institute of Mathematics and Its Applications), and SAMSI (Statistical and Applied Mathematical Sciences Institute) dealing with the theme of shape and statistics on shape manifolds, points to the likely development of new techniques to do more sophisticated statistical analysis of the ATR problem, and to the development of an ATR theory to predict optimal system performance. Also these conferences and workshops point to a wide array of applications, including important applications of shape and shape statistics to medical images, automated inspection, and image segmentation.

+ Computationally Efficient

While the description of shape and the metrics above involve some rather sophisticated mathematics, in many instances the shape coordinates and metric values are easily and directly computable via simple and fast algorithms involving minimal computational resources.

+ Maximally Robust

Because the metrics are based on shape, they are in some sense the "best possible" matching criteria as far as target geometry is concerned. As a result the metrics should be maximally robust to sensor error, pixelization, or small target variations. Our recent studies with large synthetic databases of feature sets bear this out. Additional tests have been performed by Ms. Olga Mendoza at AFRL, Wright-Patterson, AFB and by researchers at the University of Illinois.

Personnel Supported:

In addition to the principal investigator, the project has provided support for two graduate research assistants: Mr. Kevin Abbott and Ms. Jennifer Snodgrass, both graduate students in the Mathematics Department at Texas A&M University.

Ms. Snodgrass was engaged in the coding and testing of a number of algorithms and in the design of computational experiments to verify various theoretical ideas that emerged during the course of our research. Ms. Snodgrass, who received her Bachelor's degree in Applied Mathematics from Rice University, completed work on her Master's degree in May 2005.

Mr. Abbott, a Ph.D. student in Mathematics, became involved in the project as a result of a graduate course in Shape Theory offered by the P.I., Dr. Stiller, in the Spring of 2004. This course presented the results of earlier AFOSR sponsored research along with background material in differential geometry and statistical shape theory. Mr. Abbott recently completed his Ph.D. dissertation on algebro-geometric aspects of shape theory in the full perspective case. He graduated last month (August 2007) and has taken a job with Metron Corp. in Arlington, Virginia. Mr. Abbott was involved with several aspects of our collaboration with researchers at the Air Force Research Lab and accompanied the P.I. on visits to Wright Patterson Air Force Base in 2006.

Faculty: Dr. Peter F. Stiller, Prof. of Mathematics and Computer Science

Graduate Students: Jennifer Snodgrass, Kevin Abbott

Publications:

Several publications dealing with this project's results have or will appear in print shortly. A copy of the two most recent are attached to this report. The others were appended to our previous reports or are in the midst of the publication process.

D. Gregory Arnold, Olga Medoza, and Peter F. Stiller, "Image Registration via Invariant Object/Image Equations and O/I-Metrics," Algorithms for Synthetic Aperture Radar Imagery XV, SPIE Defense and Security Symposium, Orlando, FL, 3/08, to appear.

Arnold, G., Stiller, P. F., and Sturtz, K., "Geometric Methods for ATR – Invariants, Object/Image Equations, and Metrics," under revision for publication, AFRL Technical Report, 45 pages (2007).

Stiller, P. F. and Abbott, K., "Recognizing Point Configurations in Full Perspective," Electronic Imaging, Vision Geometry XV, Vol. 6499, San Jose, CA, 12 pages (2007).

Stiller, P. F. and Arnold, D. G., "Mathematical Aspects of Shape Analysis for Object Recognition," Electronic Imaging, Visual Communications and Image Processing, Vol. 6508, San Jose, CA, 12 pages (2007).

Stiller, P. F., "Robustness and statistical analysis of object/image metrics," Electronic Imaging, Vision Geometry XIV, Vol. 6066, San Jose, CA, 1/06, 9 pages (2006).

Arnold, G., Stiller, P. F., and Sturtz, K., "Object-Image Metrics for Generalized Weak Perspective Projection," chapter in Statistics and Analysis of Shapes, Editors Hamid Krim and Anthony Yezzi, Jr., Birkhauser, pp. 253-279 (2006).

Stiller, P. F., "Recognition of Configurations of Lines I – Weak Perspective Case," Proceedings SPIE Int'l Symposium on Optical Science and Technology, Mathematical Methods in Pattern and Image Analysis, Vol. 5916, Jaako Astola, Editor, San Diego, CA, 8/05, 13 pages (2005).

Stiller, P. F., "The Relationship Between Shape under Similarity Transformations and Shape under Affine Transformations," Proceedings SPIE Int'l Symposium on Optical Science and Technology, Mathematics of Data/Image Coding, Compression, and Encryption, with Applications, Vol. 5561, Mark Schmalz, Editor, Denver, CO, 8/04, pp. 108-116 (2004).

Stiller, P. F., "Vision metrics and object/image relations II: Discrimination metrics and object/image duality," Electronic Imaging, Vision Geometry XII, Vol. 5300, San Jose, CA, pp. 74-85 (2004).

Interactions/Transitions:

In June 2004, Dr. Stiller visited the Air Force Research Laboratory's Target Recognition Branch AFRL/SNAT where plans for collaborative work were made and several of the topics in the proposal were discussed.

In August 2004, Dr. Stiller attended the SPIE International Conference on Optical Science and Technology in Denver for the conference on Mathematics of Data/Image Encoding, Compression, and Encryption VI, with Applications. He presented a paper entitled "The Relationship Between Shape under Similarity Transformations and Shape under Affine Transformations." At the meeting Dr. Stiller continued discussions with Dr. Mark Schmaltz of Florida State University on possible novel applications of our research on metrics for object recognition to the completely different problem of evaluating data compression and encryption schemes.

Also in August 2004, Dr. Stiller visited Vexcel Corporation in Boulder, Colorado and presented a talk entitled "Shape Theory and Invariant Metrics for Object and Target Recognition." His visit was hosted by Dr. Carolyn Johnston. Dr. Stiller was originally put in contact with Dr. Johnston several years ago by Dr. Arje Nachman of the Air Force Office of Scientific Research.

From January 17, 2005 to January 23, 2005 Dr. Stiller again visited the Air Force Research Laboratory's Target Recognition Branch AFRL/SNAT to continue his research collaboration with Dr. Greg Arnold. Dr. Stiller returned to AFRL/SNAT in May and June of 2005. During that visit, work on the weak perspective case for point features was completed and written up in an invited book chapter entitled "Object-Image Metrics for Generalized Weak Perspective Projection" which has now appeared in a volume entitled *Statistics and Analysis of Shapes,* edited by Professor Hamid Krim of North Carolina State University.

In May 2005, Dr. Stiller attended the AFOSR Program Review at North Carolina State University hosted by Dr. Jon Sjogren, AFOSR and Professor Hamid Krim, NC State. Dr. Stiller spoke on "Shape, Shape Matching Metrics, and Learning Shape by Sampling (Shapelets)" jointly with Dr. Greg Arnold, AFRL/SNAT.

While visting AFRL's Target Recognition Branch AFRL/SNAT in June 2005, Dr. Stiller held a number of discussions with Mr. Ron Dilsavor of SET Associates, Inc. concerning ways to use this project's results to recognize objects in SAR images.

In August 2005, Dr. Stiller attended the SPIE International Conference on Optical Science and Technology in San Diego for the conference on Mathematical Methods in Pattern and Image Analysis. He presented a paper entitled " Recognition of Configurations of Lines I – Weak Perspective Case."

Dr. Stiller was an invited participant in the IMA Workshop on New Mathematics and Algorithms for 3-D Image Analysis, at the Institute for Mathematics and its Applications, University of Minnesota, Minneapolis, MN, Jan. 9-12, 2006. Several researchers from ARFL also attended, including Dr. Greg Arnold, AFRL/SNAT and Ms. Olga Mendoza, a recent hire at AFRL. During the workshop, Dr. Stiller and Dr. Arnold held discussions with Dr. Guillermo Sapiro, Department of Electrical and Computer Engineering, University of Minnesota, concerning ideas for point cloud matching and with Professor Peter Olver concerning aspects of differential invariants.

Dr. Stiller returned to the Institute for Mathematics and its Applications, University of Minnesota, in April 2006 to attend the Workshop on Shape Spaces (April 3-7, 2004) organized by Professor David Mumford. Joint with Dr. Arnold, AFRL/SNAT, Dr. Stiller held discussions with Dr. T. J. Klausutis, AFRL, Eglin, AFB, who also attended the workshop. These discussions concerned applications of shape theoretic techniques to various Air Force target recognition problems.

In May 2006, Dr. Stiller visited Dr. Arnold at the Air Force Research Laboratory's Target Recognition Branch AFRL/SNAT. The purpose was to engage in collaborative research on a number of problems including the 3D reconstruction from motion problem and various shape statistics problems. Dr. Stiller also worked with two graduate students visiting AFRL/SNAT for the summer. During this visit Dr. Arnold and Dr. Stiller traveled to Purdue University to speak with Dr. Mirelle Boutin (mentioned above) about her work on invariants for unordered point features and to give a joint talk in the Electrical Engineering Department. This resulted in Dr. Arnold and Dr. Stiller being invited by Prof. Boutin to submit a paper to the SPIE conference on Electronic Imaging, Visual Communications in Image Processing which was held during January 2007 in San Jose.

In August 2006 Dr Stiller returned to Wright Patterson AFB to again coordinate research efforts with Dr. Arnold and to attend the Multi-Modal Biometrics Workshop hosted jointly by the Human Effectiveness Biosciences and Protection Division and the Sensors ATR Division of AFRL at Wright Patterson AFB. The goal was to exchange ideas on recognition and identification technologies. It was an opportunity for us to explore applications of our recognition techniques to biometric problems such as face/body recognition and gait analysis.

On 30 January to 1 February 2007 Dr. Stiller attended SPIE's Conference on Electronic Imaging held in San Jose, CA to present two papers. The first paper "Recognizing Point Configurations in Full Perspective" was joint with his Ph.D. student Kevin Abbott and was presented in Vision Geometry XIV. Dr. Stiller chaired the session on Surface Analysis and Reconstruction in that conference. The second paper "Mathematical Aspects of Shape Analysis for Object Recognition" was an invited paper for the session on Visual Communications and Image Processing. One important research contact to come out of this meeting was a series of discussions with the 3D TV group at Phillips Electronics. They are interested in using our techniques as a tool for adding depth information to existing video content. In addition, we learned that researchers at the University of Illinois are using our Object/Image metric for the affine case in a number of computer vision experiments.

From March 3rd to March 7th 2007 Dr. Stiller participated in the workshop on New Directions in Complex Data Analysis for Emerging Applications that was held under the sponsorship of AFOSR and NSF in Breckenridge, Colorado. In addition to giving a brief presentation entitled "Algebraic Geometry, Shape, and Understanding Configurations from Projections to Lower Dimensions with Applications to Object Recognition and Image Understanding," Dr. Stiller participated in various panel discussions. While at the workshop, Dr. Stiller began discussions with Dr. Louis Scharf on a geometric approach to a long standing problem in signal processing. This problem can be reinterpreted as minimizing a distance in a Grassmannian between two subvarieties, one of which comes from the k-secants of a rational normal curve and the other of which is a standard Schubert cycle.

After the completion of the Spring semester in May 2007, Dr. Stiller made another visit to Wright Patterson AFB to again coordinate research efforts with Dr. Arnold. The focus was on updating and expanding our joint paper "Geometric Methods for ATR - Invariants, Object/Image Equations, and Metrics" for publication. In addition we continued discussions with Dr. Matt Ferrara at AFRL on 3D target reconstruction from multiple 1D radar range profiles.

On June 21st and 22nd 2007 Dr. Stiller attended the AFOSR Sensing Program Review at Harvard University. He spoke on "Shape, Shape Statistics, and Reconstruction."

Dr Stiller was an invited attendee at the SAMSI Summer Program on the Geometry and Statistics of Shape. This program ran from July 7th through July 13th, 2007 at the Statistical and Applied Mathematical Sciences Institute (SAMSI) in Research Triangle Park, NC.

In August 2007 Dr. Stiller returned to AFRL, Wright-Patterson to continue his collaboration with researchers there. In addition a new collaborative effort was begun with Ms. Olga Mendoza (AFRL/SNAT) dealing with applications of our object/image metrics in the affine case to image registration problems.

New Discoveries, Inventions, or Patent Disclosures:

Beyond the research results discussed above, there are no new discoveries, inventions, or patent disclosures.

Person completing this report:

Dr. Peter F. Stiller
Professor of Mathematics and Computer Science
Associate Director of the Institute for Scientific Computation
Phone: (409) 862-2905
Fax: (409) 845-5827
Date: 30 September 2007

Attachments:

1) Summary of our talk at the Breckenridge workshop.
2) Abstract of our paper, "Image Registration via Invariant Object/Image Equations and O/I-Metrics."
3) Copy of our slides from the Sensing Program Review at Harvard.
4) Copy of our recent paper "Recognizing Point Configurations in Full Perspective" joint with Kevin Abbott.
5) Copy of our recent paper "Mathematical Aspects of Shape Analysis for Object Recognition" joint with D. Gregory Arnold (AFRL).

Note: 3) – 5) provided as separate files in the electronic version of this report.

Attachment #1

New Directions in Complex Data Analysis for Emerging Applications

Breckenridge, Colorado
March 4-7, 2007

Summary of Talk: "Algebraic Geometry, Shape, and Understanding Configurations from Projections to Lower Dimensions with Applications to Object Recognition and Image Understanding" by Dr. Peter F. Stiller, Professor of Mathematics and Computer Science, Associate Director Institute for Scientific Computation, Texas A&M University.

Efficiently recognizing three dimensional arrangements of features on an object from a single two dimensional view requires an approach that is view and pose invariant. Existing methods often rely on computationally expensive template matching. Those methods use comparisons against templates created for all possible views; with the infinite number of possibilities being approximated by some finite number of views. To carry out an invariant approach to target recognition, we need to exploit properties and relationships that are geometrically intrinsic to the objects and/or images being compared.

Our approach to view and pose independence (as well as coordinate independence) starts with a characterization of a configuration of features by its geometric invariants. The specific group to which things should be invariant is a function of the sensor type. We then derive a fundamental set of equations that express, in an invariant way, the relationship between the 3D geometry and its "residual" in a 2D (or 1D) image. These equations completely and invariantly describe the mutual 3D/2D constraints. Once derived, they can be exploited in a number of ways. For example, from a given 2D configuration, we are able to determine a set of nonlinear constraints on the geometric invariants of the 3D configurations capable of producing that given 2D configuration, and thereby arrive at a test for determining the object being viewed. Conversely, given a 3D geometric configuration (features on an object), we are able to find a set of equations that constrain the invariants of the images of that object; helping to determine if that object appears in selected images. With these results in hand, we plan in future work to focus on three major problems: 1) object/image metrics on shape spaces to provide a distance (difference) between two object configurations, two image configurations, or an object and an image pair in pose invariant, coordinate free terms, 2) reconstruction of an object's 3D shape from 2D sensed information, either from multiple sensors or multiple images of a moving object, 3) statistical issues surrounding random shapes, distributions of shapes, and noise in object recognition.

Issues and Collaborations Arising From the Conference: One topic that arose in several of the presentations, was the issue of dealing with data on certain manifolds, most notably Grassmann manifolds. In the work of Peterson, Kirby and in my own work complex image data is represented by data points in a Grassmannian. Appropriate metrics and also procedures for fitting subvarieties to such data need to be developed. The general question of invariant features of high dimensional data under projections to lower dimensions is also an interesting one. It appears that some aspects of our techniques could be applied to such high dimensional problems. Finally, an interesting signal processing problem, introduced to the author by Louis Scharf at the meeting, appears to have a nice geometric formulation in terms of secant varieties of rational normal curves, where the same sort of metrics on Grassmannians play a role in finding the optimal answer. We are currently investigating this.

Attachment #2

Image Registration via Invariant Object/Image Equations and O/I-Metrics

By D. Gregory Arnold, Olga Mendoza, and Peter F. Stiller

The problem of single-view recognition is central to many target recognition and computer vision tasks. Understanding how information available in a single image of an object or scene, be it an optical image, a SAR image, or a radar range profile, relates to the target object's or scene's geometry is a key step in building reliable identification algorithms. Likewise such knowledge is critical to understanding how two different images of the same object or scene are related. For example, without a priori knowledge of a sensor's viewpoint, an object's pose, or a sensor's parameters, it is difficult to efficiently recognize a three-dimensional arrangement of features (such as a geometric configuration of lines and/or points) on an object or to efficiently register two images of the same object or scene. What is needed is an approach that is invariant to changing viewpoints, adjustments in the sensor parameters, or variations in the pose.

In recent work the authors have developed such an approach to object recognition, and the goal of this paper is to apply the same techniques to the registration problem. To carry out their recognition work, they started with a characterization of a configuration of features by its geometric invariants. The specific transformation group to which things needed to be invariant was a function of the sensor type. They then derived a fundamental set of equations that expressed, in an invariant way, the relationship between the 3D geometry and its "residual" in a 2D image. These equations completely and invariantly described the mutual 3D/2D constraints. Once derived, the equations could be exploited in a number of ways. For example, from a given 2D configuration, they could determine a set of nonlinear constraints on the geometric invariants of the 3D configurations capable of producing that given 2D configuration, and thereby arrive at a test for determining the object being viewed. Here having two images of the same 3D configuration would add additional constraints and tell you a fair amount about the relationship between the two images – thereby assisting with the registration of those images. That is something we take up in this paper. Conversely, given a 3D geometric configuration (features on an object), a set of equations that constrain the invariants of the images of that object were derived; helping to determine if that object appears in selected images. These equations also play a role in registration of different images of the same scene or object. They give us an understanding of the locus of all images and the flow from image to image as the sensor moves. We discuss applications of this in the paper. Finally, the authors have developed certain natural invariant metrics (called OI-metrics) on the relevant shape spaces. Thes metrics provide a distance (difference) measure between two object configurations or two image configurations and express the distance (failure to match) between, say, an image-image pair. These metrics are pose and view invariant and can be expressed in coordinate free terms.

For example, consider the generalized weak perspective model of image formation, which is appropriate to optical images when the object or scene in the far field. Here the relevant invariance is to the affine group of transformations. In this case the OI-metric for images will measure the failure of two images to differ by an affine transformation. As such, it provides a quantification of the drift phenomenon seen in image registration done via affine mappings.

By understanding the contribution of a single image toward the recognition or recovery of the geometry/shape of the object or scene for different sensors, it will be easier to develop methods to integrate information from multiple images taken by uncalibrated, distributed sensors of varying types, or to make use of a series of images taken by a single sensor of a moving object. We investigate in this paper how to apply our invariant techniques to the problem of registering those multiple images.

Shape, Shape Statistics, and Reconstruction

Peter F. Stiller

Texas A&M University

21 June 2007

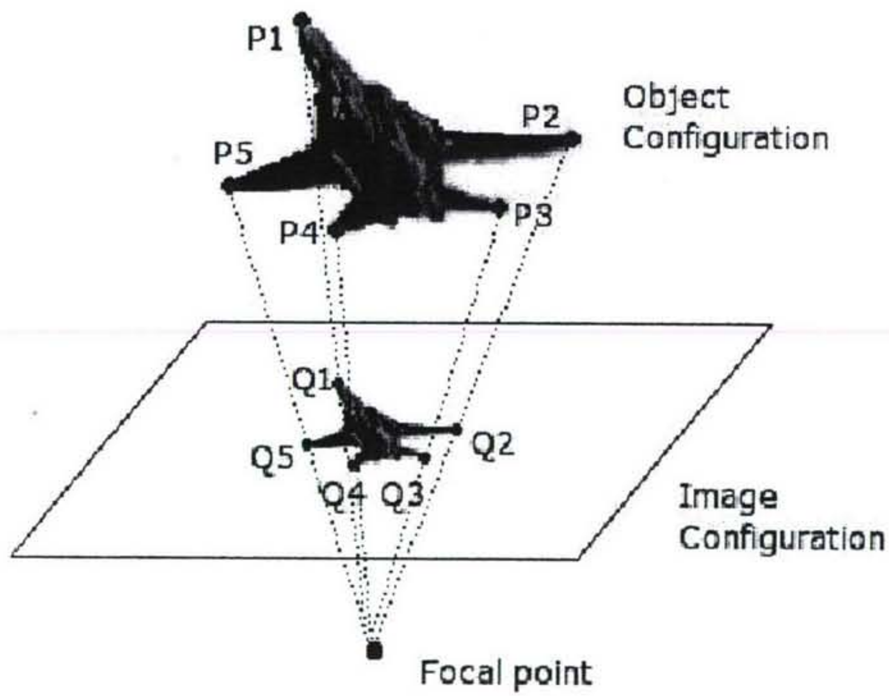
Motivation

- "How can we get a computer to determine what object produced a given image?"

- GOAL: Develop mathematical methods for Automated Target Recognition (ATR) that do not depend on
  - the object's position in relation to the sensor.
  - internal sensor parameters (such as focal length).
  - the choice of coordinate system used to represent the object and/or image.

Basic Idea

- We represent an object by an k-tuple of points in 3-space (object configuration).
- Taking a picture of the object projects the object configuration onto a plane (image configuration).



- To achieve the desired invariance, we want to match *shapes* of object configurations with *shapes* of image configurations under appropriate projections.

- The *shape* of a configuration is its equivalence class under the action of some group of transformations.

- Let G be a group of transformations on $\mathbb{R}^n$. $P_1, \ldots, P_k$ and $P'_1, \ldots, P'_k$ have the same shape with respect to G if there is some $g \in G$ such that $g(P_i) = P'_i$.

- Observations:
  - Our configurations are *ordered* k-tuples of points.
  - If two configurations have the same shape, they must have the same number of points.
  - An image must consist of the same number of points as its corresponding object.

The Generalized Weak Perspective Model

- Points are considered in affine space ($\mathbb{R}^2$ or $\mathbb{R}^3$).

$$P = \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix} \quad \text{or} \quad Q = \begin{pmatrix} u \\ v \\ 1 \end{pmatrix}$$

- We consider generalized weak perspective (GWP) projections

$$T = \begin{pmatrix} t_{11} & t_{12} & t_{13} & t_{14} \\ t_{21} & t_{22} & t_{23} & t_{24} \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

where T has rank 3.

- Two configurations have the same shape if they differ by an affine transformation

$$\left(\begin{array}{cc|c} a_{11} & a_{12} & r_1 \\ a_{21} & a_{22} & r_2 \\ \hline 0 & 0 & 1 \end{array}\right) \quad \text{or} \quad \left(\begin{array}{ccc|c} b_{11} & b_{12} & b_{13} & s_1 \\ b_{21} & b_{22} & b_{23} & s_2 \\ b_{31} & b_{32} & b_{33} & s_3 \\ \hline 0 & 0 & 0 & 1 \end{array}\right).$$

The Affine Shape Spaces

- The space of object shapes - $\left(\mathbb{R}^3\right)^n / Aff(3)$
  The space of image shapes - $\left(\mathbb{R}^2\right)^n / Aff(2)$.
- Represent a configuration (k points in $\mathbb{R}^n$) as a matrix

$$M = \begin{pmatrix} x_{11} & x_{21} & & x_{k1} \\ x_{12} & x_{22} & \cdots & x_{k2} \\ \vdots & \vdots & & \vdots \\ x_{1n} & x_{n2} & & x_{kn} \\ 1 & 1 & & 1 \end{pmatrix}$$

- Must assume that the points of the configuration do not all lie in a single hyperplane.
- Associate to the config. $K^{k-n-1} \subset \mathbb{R}^k$ (the null-space of M).

- dim $\left(K^{k-n-1}\right) = k - n - 1$ since the config. is non-coplanar.

- K^{k-n-1} is invariant under the left action of $Aff(n)$

- Every K^{k-n-1} is contained in

$$H^{k-1} = \left\{ (v_1, \ldots, v_k) \in \mathbb{R}^k \middle| \sum_{i=1}^{k} v_i = 0 \right\}$$

- The **Affine Shape Space** of configurations of k points in $\mathbb{R}^n$ is the Grassmannian $Gr(k-n-1, H^{k-1}) \subset Gr(k-n-1, k)$.

$$P_1, P_2, \ldots, P_k \quad \longleftrightarrow \quad \left[K^{k-n-1}\right] \in Gr(k-n-1, H^{k-1}).$$

- Affine Object Space ($n = 3$)

$$\mathcal{O}_k = Gr(k-4, H^{k-1}) \subset Gr(k-4, k)$$

- Affine Image Space ($n = 2$)

$$\mathcal{I}_k = Gr(k-3, H^{k-1}) \subset Gr(k-3, k)$$

Shape Coordinates

- Embed $Gr(k-n-1,k)$ into $\mathbb{P}^{\binom{k}{k-n-1}-1} \cong \mathbb{P}^{\binom{k}{n+1}-1}$ as a projective variety by the Plücker embedding.
  - Embeds $Gr\left(k-n-1, H^{k-1}\right)$ as a subvariety.
- For an affine shape K, choose a representative config. $P_1, \ldots, P_n$ and write as a matrix.

$$M = \begin{pmatrix} x_{11} & x_{21} & & x_{k1} \\ x_{12} & x_{22} & \cdots & x_{k2} \\ \vdots & \vdots & & \vdots \\ x_{1n} & x_{n2} & & x_{kn} \\ 1 & 1 & & 1 \end{pmatrix}$$

- For $1 \leq i_1 < i_2 < \ldots < i_{n+1} \leq k$

$$m_{i_1 i_2 \ldots i_{n+1}} = \det \begin{pmatrix} x_{1i_1} & x_{1i_2} & \cdots & x_{1i_{n+1}} \\ x_{2i_1} & x_{2i_2} & \cdots & x_{2,i_{n+1}} \\ \vdots & \vdots & \ddots & \vdots \\ x_{ni_1} & x_{ni_2} & \cdots & x_{ni_{n+1}} \\ 1 & 1 & \ldots & 1 \end{pmatrix}$$

- The map from our affine shape space $Gr(k-n-1, H^{k-1})$ to $\mathbb{P}^{\binom{k}{n+1}-1}$ is given by

$$K^{k-n-1} \longmapsto \left(m_{12\ldots n+1} : \ \ldots \ : m_{k-n\ldots k}\right)$$

(all minors)

- The homogeneous coordinates

$$\left(m_{12\ldots n+1} : \ \ldots \ : m_{k-n\ldots k}\right)$$

are called the **shape coordinates** of the configuration $P_1, \ldots, P_n$.

The Object/Image Relations

- Want necessary and sufficient conditions for an image configuration $Q_1, \ldots, Q_k$ to be a GWP projection of an object configuration $P_1, \ldots, P_k$.

- The locus of matching object/image pairs should be should be a subvariety

$$V \subset \mathcal{O}_k \times \mathcal{I}_k \subset \mathbb{P}^{\binom{k}{4}-1} \times \mathbb{P}^{\binom{k}{3}-1}$$

- The relations should be a system of bihomgeneous polynomials in the object and image shape coordinates whose zero locus is V.

Theorem 1. *Let $P_1, \ldots, P_k$ be an object configuration with shape coordinates $(\ldots : m_{i_1 i_2 i_3 i_4} : \ldots)$ and let $Q_1, \ldots, Q_k$ be an image configuration with shape coordinates $(\ldots : n_{i_1 i_2 i_3} : \ldots)$. Then $Q_1, \ldots, Q_k$ is a GWP projection of $P_1, \ldots, P_k$ if and only if*

$$\sum_{1 \leq \lambda_1 < \lambda_2 \leq k} m_{\alpha_1,\alpha_2,\lambda_1,\lambda_2} n*_{\lambda_1,\lambda_2,\beta_1,\ldots,\beta_{k-5}} = 0$$

for all choices of $1 \leq \alpha_1 < \alpha_2 \leq k$ and $1 \leq \beta_1 < \beta_2 < \ldots < \beta_{k-5} \leq k$ The expressions $m_{\alpha_1,\alpha_2,\lambda_1,\lambda_2}$ and $n*_{\lambda_1,\lambda_2,\beta_1,\ldots,\beta_{k-5}}$ should be treate as skew-symmetric in the entries of the indices.

- In terms of the standard shape coordinates

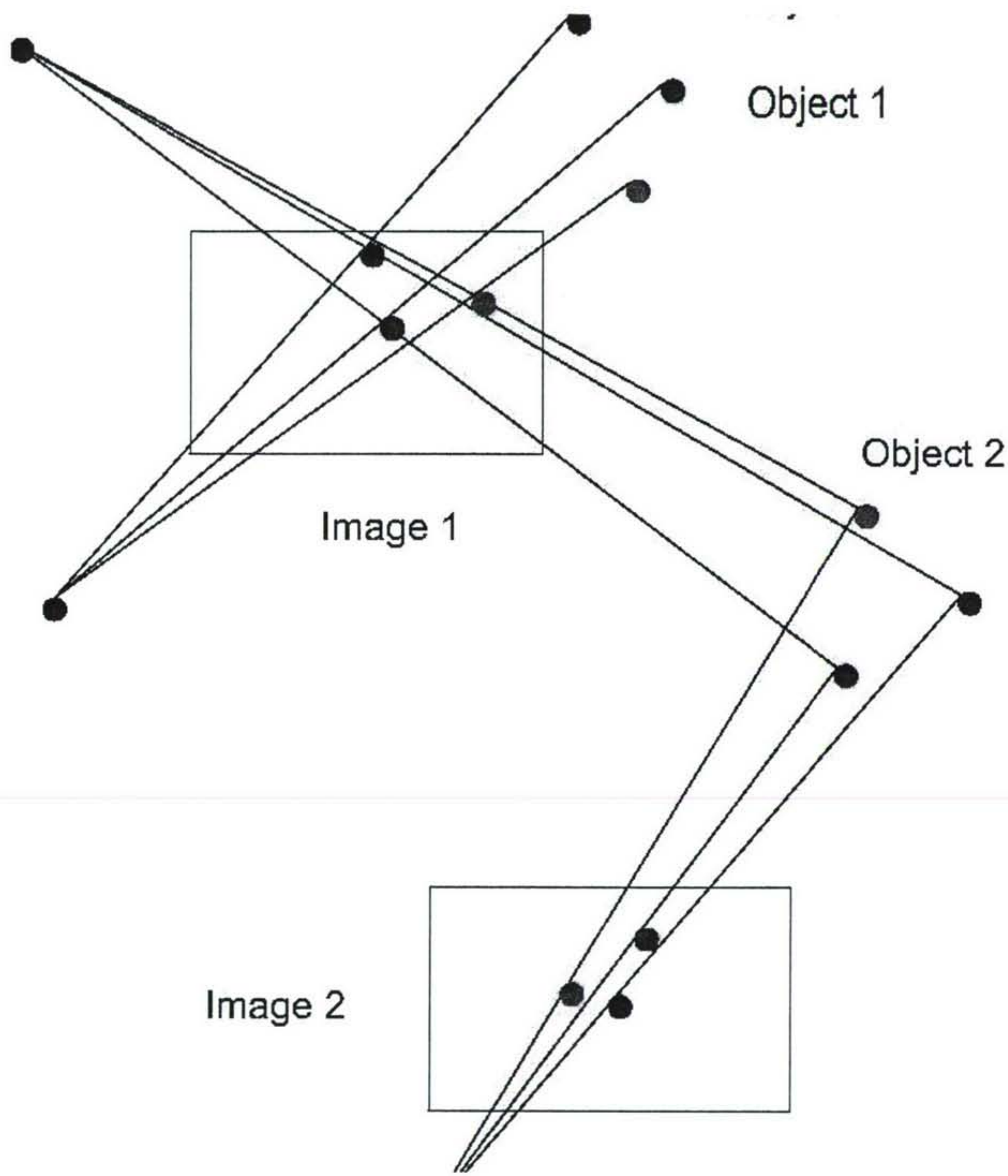
$$\sum_{1 \leq \lambda_1 < \lambda_2 \leq k} \epsilon_{\lambda_1,\lambda_2} m_{\alpha_1,\alpha_2,\lambda_1,\lambda_2} n_{\gamma_1,\gamma_2,\gamma_3} = 0$$

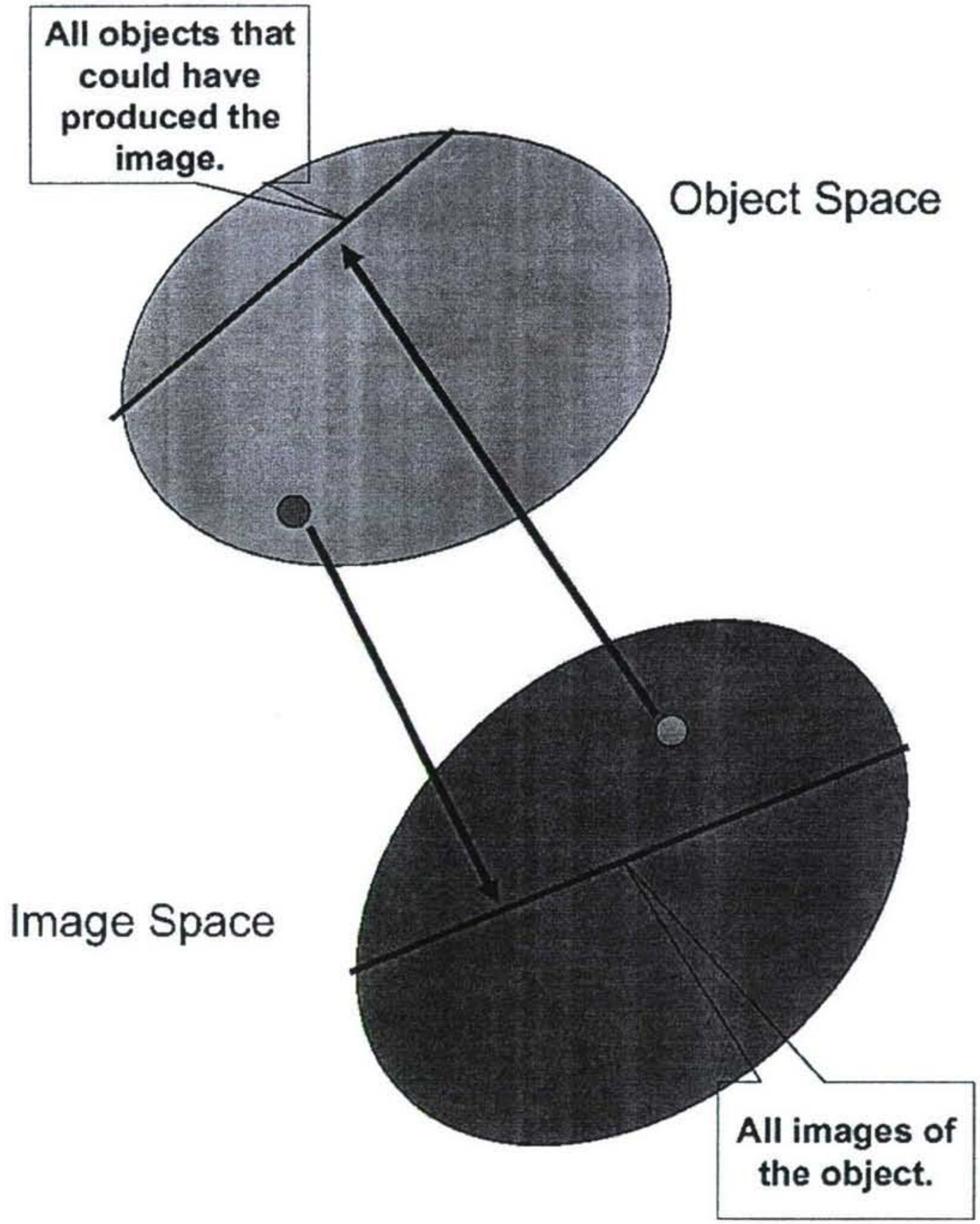
  - $\{\gamma_1, \gamma_2, \gamma_3\}$ is the complement of $\{\lambda_1, \lambda_2, \beta_1, \ldots, \beta_{k-5}\}$

- $\epsilon_{\lambda_1,\lambda_2}$ is the sign of the permutation $\gamma_1, \gamma_2, \gamma_3, \lambda_1, \lambda_2, \beta_1, \ldots, \beta_{k-5}$

Example 2. *5 points*

$$m_{1234}n_{124} - m_{1235}n_{124} + m_{1245}n_{123} = 0$$
$$m_{1234}n_{135} - m_{1235}n_{134} + m_{1345}n_{123} = 0$$
$$m_{1234}n_{145} - m_{1245}n_{134} + m_{1345}n_{124} = 0$$
$$m_{1235}n_{145} - m_{1245}n_{135} + m_{1345}n_{125} = 0$$
$$m_{1234}n_{235} - m_{1235}n_{234} + m_{2345}n_{123} = 0$$
$$m_{1234}n_{245} - m_{1245}n_{234} + m_{2345}n_{124} = 0$$
$$m_{1235}n_{245} - m_{1245}n_{235} + m_{2345}n_{134} = 0$$
$$m_{1234}n_{345} - m_{1345}n_{234} + m_{2345}n_{134} = 0$$
$$m_{1235}n_{345} - m_{1345}n_{235} + m_{2345}n_{135} = 0$$
$$m_{1245}n_{345} - m_{1345}n_{245} + m_{2345}n_{145} = 0$$


Object 1
Object 2
Image 1
Image 2


All objects that could have produced the image.
Object Space
Image Space
All images of the object.

Metrics

- Let K^{k-4} and $\tilde{K}^{k-4}$ be two object shapes.
- Choose orthonormal bases for them.
- Arrange those vectors as the columns of two orthonormal matrices K and $\tilde{K}$ (each $k \times (k-4)$).
- Compute the singular values of $K^T\tilde{K}$, $\sigma_1, \ldots, \sigma_{k-4}$.
- $\theta_i = \arccos \sigma_i$ — the principal angles between K^{k-4} and $\tilde{K}^{k-4}$
- Distance in affine object space is

$$d_{Obj}(K^{k-4}, \tilde{K}^{k-4}) = \sqrt{\sum_{i=1}^{k-4} \theta_i^2}$$

- Distance in affine image space is defined similarly

$$d_{Img}(L^{k-3}, \tilde{L}^{k-3}) = \sqrt{\sum_{i=1}^{k-3} \phi_i^2}$$

We may compute a "distance" between an object shape K^{k-4} and an image shape L^{k-3} in 3 ways:

- Working in object space

$$d^1_{O/I}(K^{k-4}, L^{k-3}) = \min_{\tilde{K}^{k-4} \subset L^{k-3}} d_{Obj}(K^{k-4}, \tilde{K}^{k-4})$$

- Working in image space

$$d^2_{O/I}(K^{k-4}, L^{k-3}) = \min_{\tilde{L}^{k-3} \supset K^{k-4}} d_{Img}(L^{k-3}, \tilde{L}^{k-3})$$

- By computing the principal angles between L^{k-3} and K^{k-4}.
  - Compute the singular values of $L^T K$, $\sigma_1, \ldots, \sigma_{k-4}$
  - $\theta_i = \arccos(\sigma_i)$

$$d_{O/I}(K^{k-4}, L^{k-3}) = \sqrt{\sum_{i=1}^{k-4} \theta_i^2}$$

Object/Image Metric Duality

Theorem 1. *For an object shape K^{k-4} and an image shape L^{k-3}*

$$d^1_{O/I}(K^{k-4}, L^{k-3}) = d^2_{O/I}(K^{k-4}, L^{k-3}) = d_{O/I}(K^{k-4}, L^{k-3})$$

Full Perspective

For a configuration $P_1, \ldots, P_k \in \mathbb{P}^n$

- We have a map $\overline{\Phi} : (\mathbb{R}^*)^k / \mathbb{R}^* \to G(n+1, k)$ given by

  $$\overline{\Phi}(a_1, \ldots, a_k) = (a_{I_1} m_{I_1} : \ldots : a_{I_N} m_{I_N})$$

  where $I_1, \ldots, I_N$ are the $(n+1)$-subsets of $\{1, \ldots, k\}$ and $a_{I_r} = \prod_{j \in I_r} a_j$.

- The **shape variety** of the configuration is $\overline{Im(\overline{\Phi})} = \overline{\mathcal{V}(P_1, \ldots, P_k)}$

- Shapes of configurations of k points are in 1-1 correspondence with varieties $\overline{\mathcal{V}}(P_1, \ldots, P_k)$

$$
\begin{aligned}
&n_{125}n_{136}n_{234}m_{1236}m_{1246}m_{1345}m_{2345}\\
&-n_{123}n_{136}n_{234}m_{1236}m_{1246}m_{1345}m_{2345}\\
&-n_{126}n_{135}n_{234}m_{1236}m_{1245}m_{1346}m_{2345}\\
&+n_{124}n_{135}n_{236}m_{1236}m_{1245}m_{1346}m_{2345}\\
&+n_{125}n_{134}n_{236}m_{1235}m_{1245}m_{1346}m_{2345}\\
&-n_{124}n_{135}n_{236}m_{1235}m_{1246}m_{1346}m_{2345}\\
&+n_{126}n_{134}n_{235}m_{1236}m_{1245}m_{1345}m_{2346}\\
&-n_{124}n_{136}n_{235}m_{1236}m_{1245}m_{1345}m_{2346}\\
&-n_{125}n_{136}n_{234}m_{1235}m_{1246}m_{1345}m_{2346}\\
&+n_{124}n_{136}n_{235}m_{1235}m_{1246}m_{1345}m_{2346}\\
&+n_{126}n_{135}n_{234}m_{1235}m_{1245}m_{1346}m_{2346}\\
&-n_{126}n_{134}n_{235}m_{1235}m_{1245}m_{1346}m_{2346} = 0.
\end{aligned}
$$

Current Work

- Shape Statistics
- Find distribution of image shapes given a distribution of object shapes and vice versa.
- Problem of non-compact space of transformations
- A few known results in the conformal case.
- Shape distributions on Grassmannians

- Shape Reconstruction
- Reconstruct object shape from multiple image shapes acquired from multiple sensors or multiple looks from a single sensor.
- Data fitting issues - what is the right error metric.
- Fitting data on Grassmannians.

Recognizing Point Configurations in Full Perspective

Kevin Abbott and Peter F. Stiller

Department of Mathematics, Texas A&M University, MS3368

College Station, Texas 77843-3368

ABSTRACT

In this paper we examine two fundamental problems related to object recognition for point features under full perspective projection. The first problem involves the geometric constraints (object-image equations) that must hold between a set of object feature points (object configuration) and any image of those points under a full perspective projection, which is just a pinhole camera model for image formation. These constraints are formulated in an invariant way, so that object pose, image orientation, or the choice of coordinates used to express the feature point locations either on the object or in the image are irrelevant. These constraints turn out to be expressions in the shape coordinates calculated from the feature point coordinates. The second problem concerns the notion of shape and a description of the resulting shape spaces. These spaces aquire certain natural metrics, but the metrics are often hard to compute. We will discuss certain cases where the computations are managable, but will leave the general case to a future paper.

Taken all together, the results in this paper provide a way to understand the relationship that exists between 3D geometry and its "residual" in a 2D image. This relationship is completely characterized (for a particular combination of features) by the above set of fundamental equations in the 3D and 2D shape coordinates. The equations can be used to test for the geometric consistency between an object and an image. For example, by fixing point features on a known object, we get constraints on the 2D shape coordinates of possible images of those features. Conversely, if we have specific 2D features in an image, we will get constraints on the 3D shape coordinates of objects with feature points capable of producing that image. This yields a test for which object is being viewed. The object-image equations are thus a fundamental tool for attacking identification/recognition problems in computer vision and automatic target recognition applications.

Keywords: object recognition, full perspective, object-image equations, shape, shape coordinates.

1. A REVIEW OF THE AFFINE CASE

We consider r points in space, which we think of as feature points on some object. We refer to this set of points as an *object configuration.* Next, we "take a picture" of the object by choosing a plane and projecting these feature points into that plane. We will call this set of points in the plane an *image configuration.*

In this section we will (1) identify the space of shapes, which are configurations modulo the action of a certain group of transformations on $\mathbb{R}^n$, $n = 2, 3$, and give global coordinates on the shape space, (2) give necessary and sufficient conditions for an image configuration to be a projection of an object configuration, and (3) define a natural metric on the shape spaces. For additional details see Arnold, Stiller, and Sturtz [1].

1.1. The Generalized Weak Perspective Projection

The type of projections we will consider are called *generalized weak perspective projections.* If we represent points in $\mathbb{R}^n$ ($n = 2$ or 3) in the form

$$P = \begin{pmatrix} x_1 \\ \vdots \\ x_n \\ 1 \end{pmatrix}$$

Further author information: (Send correspondence to P.F.S.)

P.F.S.: E-mail: stiller@isc.tamu.edu; Telephone: 979-862-2905; Fax: 979-845-5827

these projections as maps from $\mathbb{R}^3$ to $\mathbb{R}^2$ take the form

$$T = \begin{pmatrix} t_{11} & t_{12} & t_{13} & t_{14} \\ t_{21} & t_{22} & t_{23} & t_{14} \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

where T has rank 3.

Now let A be an invertible 3×3 matrix and let B be an invertible 4×4 matrix. It turns out that if T is a generalized weak perspective projection, then ATB is a generalized weak perspective projection if an only if A and B are *affine transformations* i.e. A and B take the form

$$\begin{pmatrix} & & & c_1 \\ & S & & \vdots \\ & & & c_n \\ 0 & \cdots & 0 & 1 \end{pmatrix}$$

where $S \in GL(n)$ and $c_1, \ldots, c_n \in \mathbb{R}$.

What does this mean in terms of our object and image configurations? Suppose $Q \in \mathbb{R}^2$ is the image of a point $P \in \mathbb{R}^3$ under a generalized weak perspective projection T, i.e. $Q = TP$. Then if we move P by some affine transformation B to another point P' and if we move Q to another point Q' by an affine transformation A, we will have that $Q' = ATB^{-1}P'$. As a result, we see that Q' is the image of P' under the generalized weak perspective projection ATB^{-1} (since A and B^{-1} are both affine transformations).

This observation shows us that by choosing to consider generalized weak perspective projections, the best that we can hope to do is relate object configurations to image configurations up to affine transformations.

1.2. The Affine Shape Spaces

As the preceding observation suggests, we should consider two configurations (object or image) equivalent if they differ by an affine transformation. In a sense equivalent configurations are the same object or image just rotated, translated, scaled, or otherwise moved by an affine transformation. We would like to construct the space of configurations of r points in $\mathbb{R}^n$ modulo the action of the group of affine transformations. These spaces would then represent the distinct objects and images independent of pose or view. To do this we must assume that the points in our configuration are non-coplanar for $n = 3$ or non-collinear for $n = 2$, which is reasonable since a configuration of coplanar points in $\mathbb{R}^3$ would in fact be a configuration of points in $\mathbb{R}^2$ and would not represent a real 3D object, etc.

Let $P_i = (x_{i,1}, \ldots, x_{i,n})$ for $i = 1 \ldots r$, $r \geq n+2$ be a configuration of r non-coplanar (or non-collinear) points in $\mathbb{R}^n$, $n = 3$ (or 2), and consider the matrix

$$M = \begin{pmatrix} x_{1,1} & x_{2,1} & & x_{r,1} \\ x_{1,2} & x_{2,2} & \cdots & x_{r,2} \\ \vdots & \vdots & & \vdots \\ x_{1,n} & x_{n,2} & & x_{r,n} \\ 1 & 1 & & 1 \end{pmatrix}$$

Now to the configuration $P_1, \ldots, P_r$ we associate an $(r-n-1)$-dimensional linear subspace, $K^{r-n-1} \subset \mathbb{R}^r$. In particular, K^{r-n-1} is the null space of M when we view M as a linear map from $\mathbb{R}^r$ to $\mathbb{R}^{n+1}$. The fact that K^{r-n-1} has dimension $r-n-1$ follows from the observation that M has rank $n+1$ as a linear map because at least one $(n+1) \times (n+1)$ minor of M has non-zero determinant due to non-coplanarity (or non-collinearity).

The important thing to notice is that if we apply an affine transformation A to our configuration we obtain a new $(n+1)\times r$ matrix

$$M' = \begin{pmatrix} x'_{1,1} & x'_{2,1} & & x'_{r,1} \\ x'_{1,2} & x'_{2,2} & \cdots & x'_{r,2} \\ \vdots & \vdots & & \vdots \\ x'_{1,n} & x'_{n,2} & & x'_{r,n} \\ 1 & 1 & & 1 \end{pmatrix} = AM$$

but the null space of M' is exactly K^{r-n-1}, the null space of M. Moreover, since $K^{r-n-1} \subset H^{r-1} = \{(v_1,\ldots,v_r) \in \mathbb{R}^r | \sum_{i=1}^{r} v_i = 0\}$, we may assign to our configuration the unique point $[K^{r-n-1}] \in G(r-n-1, H^{r-1})$, the Grassmannian of $(r-n-1)$-dimensional subspaces in the $(r-1)$-dimensional space H^{r-1}, a well understood compact manifold of dimension $n(r-n-1)$.

DEFINITION 1.1. *We call the manifold $X = G(r-n-1, H^{r-1})$ the affine shape space for configurations of r points in $\mathbb{R}^n$. If $n=3$, we will call $X = G(r-4, r-1)$ affine object space (or just object space) and refer to points in this space as object shapes. If $n=2$, we will call $X = G(r-3, r-1)$ affine image space (or just image space) and refer to its elements as image shapes.*

Every point in X is of the form $[K^{r-n-1}]$ for some configuration $P_1,\ldots,P_r \in \mathbb{R}^n$, and most importantly, if two configurations $P_1,\ldots,P_r \in \mathbb{R}^n$ and $P'_1,\ldots,P'_r \in \mathbb{R}^n$ give the same point in X, then they differ by an affine transformation.

1.3. The Plücker Embedding

Since X is a real manifold, we can find local coordinates for a point $[K] \in X$; however, since we ultimately want to give relations that tell us when an image configuration is a projection of an object configuration, it would be more convenient to find global coordinates on X. We may do so by mapping X into a projective space via the *Plücker embedding*. In general, the Plücker embedding embeds a Grassmannian $G(n,r)$ (n-dimensional subspaces of an r-dimensional vector space, V^r) in the projective space $\mathbb{P}\left(\bigwedge^{r-n} V^r\right) \cong \mathbb{P}^{\binom{r}{r-n}-1} \cong \mathbb{P}^{\binom{r}{n}-1}$ as a projective variety in the following way: let $[K] \in G(n,r)$. Then K is the intersection of $r-n$ hyperplanes in our vector space V^r, where each hyperplane is given by a linear form

$$0 = \sum_{i=1}^{r} k_{j,i}\widehat{e_i}, \quad j = 1,\ldots,r-n$$

where $e_1,\ldots,e_r$ is a basis for V^r and $\widehat{e_i}$ are the dual basis. More simply put, K is the null space of the matrix

$$L = \begin{pmatrix} k_{1,1} & k_{1,2} & \cdots & k_{1,r} \\ k_{2,1} & k_{2,2} & \cdots & k_{2,r} \\ \vdots & \vdots & \ddots & \vdots \\ k_{r-n,1} & k_{r-n,2} & \cdots & k_{r-n,r} \end{pmatrix}$$

Now for each $1 \leq i_1 < i_2 < \ldots < i_{r-n} \leq r$ we define $[i_1, i_2, \ldots, i_{r-n}]$ to be the determinant of the $(r-n)\times(r-n)$ minor of L whose columns are the $i_1, i_2, \ldots, i_{r-n}$ columns of L, i.e.

$$[i_1, i_2, \ldots, i_{r-n}] = \det \begin{pmatrix} k_{1,i_1} & k_{1,i_2} & \cdots & k_{1,i_{r-n}} \\ k_{2,i_1} & k_{2,i_2} & \cdots & k_{2,i_{r-n}} \\ \vdots & \vdots & \ddots & \vdots \\ k_{r-n,i_1} & k_{r-n,i_2} & \cdots & k_{r-n,i_{r-n}} \end{pmatrix}$$

The Plücker embedding is now defined to be the map

$$\begin{array}{llll} \Phi_{n,r}: & G(n,r) & \longrightarrow & \mathbb{P}^{\binom{r}{n}-1} \\ & [K] & \longmapsto & ([1,2,\ldots,r-n] : \ldots : [n+1, n+2, \ldots, r]) \qquad \text{(all minors)} \end{array}$$

and the homogeneous coordinates of $\Phi_{n,r}([K])$ are called the *Plücker coordinates* of K.

It is important to note that this map does not depend on our choice of hyperplanes, but does depend on our choice of basis for V^r. We should also note that this map does in fact embed $G(n,r)$ as a closed projective variety in $\mathbb{P}^{\binom{r}{n}-1}$. In other words, $\Phi_{n,r}(G(n,r))$ is the zero locus of some system of homogeneous polynomials $f_1,\ldots,f_s$ in the variables $x_{1,2,\ldots,r-n};\ldots;x_{n+1,\ldots,r}$ with coefficients in the base field of V^r. We use the variables $x_{1,2,\ldots,r-n};\ldots;x_{n+1,\ldots,r}$ to indicate that the $x_{i_1,i_2,\ldots,i_{r-n}}$ coordinate of $\Phi_{n,r}([K])$ is $[i_1,\ldots,i_{r-n}]$. The equations $f_i=0,\ 1\le i\le s$ are known as the Plücker relations (see [3] or [4].

One way to give global coordinates on $X=G(r-n-1,H^{r-1})$, would be to embed X into the projective space $\mathbb{P}_{\mathbb{R}}^{\binom{r-1}{n}-1}$ via the Plücker embedding $\Phi_{r-n-1,r-1}$. However, this would require us to choose a basis for H^{r-1}. Fortunately, there is a very natural way to avoid this problem.

Since $K^{k-n-1}\subset H^{r-1}\subset\mathbb{R}^r$ we may view X as a submanifold of $G(r-n-1,r)$, in which case $\Phi_{r-n-1,r}$ embeds X in $\mathbb{P}_{\mathbb{R}}^{\binom{r}{n+1}-1}$ as a subvariety of $\Phi_{r-n-1,r}(G(r-n-1,r))$. Under this map, a configuration $P_i=(x_{i,1},\ldots,x_{i,n}),\ i=1,\ldots,r$ is mapped into $\mathbb{P}_{\mathbb{R}}^{\binom{r}{n+1}-1}$ by taking all the determinants of the maximal minors of our original feature point matrix

$$M=\begin{pmatrix} x_{1,1} & x_{2,1} & & x_{r,1}\\ x_{1,2} & x_{2,2} & \cdots & x_{r,2}\\ \vdots & \vdots & & \vdots\\ x_{1,n} & x_{n,2} & & x_{r,n}\\ 1 & 1 & & 1\end{pmatrix}$$

Embedding our shape space X into $\mathbb{P}_{\mathbb{R}}^{\binom{r}{n+1}-1}$ in this fashion is in some sense a more natural way to give global coordinates on X than embedding it into $\mathbb{P}_{\mathbb{R}}^{\binom{r-1}{n}-1}$. This method allows us to work directly with the matrix determined by our configuration rather than forcing us to choose a basis for H^{r-1} and then rewriting our basis for K^{r-n-1} in terms of our chosen basis for H. Also, as we will see later in this paper, this method is more closely related to the one that we will use in the full perspective case.

DEFINITION 1.2. *Given a configuration $P_1,\ldots,P_r\in\mathbb{R}^n$ we will refer to the Plücker coordinates of K^{r-n-1} viewed as a subspace of $\mathbb{R}^r$ (rather than H^{k-1}) as the shape coordinates of the configuration $P_1,\ldots,P_r$.*

1.4. The Object/Image Relations

Given an object configuration $P_1,\ldots,P_r$ and an image configuration $Q_1,\ldots,Q_r$ we want to give necessary and sufficient conditions (the object-image relations) for the Q_i to be a generalized weak perspective projection of the P_i. Recall that we view our object space X as a subvariety of $\mathbb{P}^{\binom{r}{4}-1}$ and our image space Y as a subvariety of $\mathbb{P}^{\binom{r}{3}-1}$. As such, we want to view the set V of pairs (K,L) where L is an image shape that comes from a generalized weak perspective projection of the object shape K (the so-called set of matching object-image pairs) as a subvariety $V\subset X\times Y\subset\mathbb{P}^{\binom{r}{4}-1}\times\mathbb{P}^{\binom{r}{3}-1}$. Therefore, our object-image relations should be a system of bihomogeneous polynomials in the object and image shape coordinates whose zero locus is precisely V.

Recall that our object shapes are linear subspaces $K^{r-4}\subset\mathbb{R}^r$ of dimension $r-4$ and our image shapes are linear subspaces $L^{r-3}\subset\mathbb{R}^n$ of dimension $r-3$. The following relates object and image shapes under generalized weak perspective projection.

THEOREM 1.3. *Let $P_1,\ldots,P_r$ be an object configuration with corresponding object shape K^{r-4} and let $Q_1,\ldots,Q_r$ be an image configuration with corresponding image shape L^{r-3}. Then the Q_i are a generalized weak perspective projection of the P_i if and only if*

$$K^{r-4}\subset L^{r-3}\subset H^{r-1}\subset\mathbb{R}^r$$

This fact and the incidence relations given in Theorem I, §5, Chapter VII of Hodge and Pedoe [4] give us our object-image relations.

THEOREM 1.4. *Let $P_i = (x_i, y_i, z_i)$, $1 \le i \le r$ be an object configuration with corresponding matrix*

$$M = \begin{pmatrix} x_1 & x_2 & & x_r \\ y_1 & y_2 & \cdots & y_r \\ z_1 & z_2 & & z_r \\ 1 & 1 & & 1 \end{pmatrix}$$

and let $Q_i = (u_i, v_i)$, $1 \le i \le r$ be an image configuration with corresponding matrix

$$N = \begin{pmatrix} u_1 & u_2 & & u_r \\ v_1 & v_2 & \cdots & v_r \\ 1 & 1 & & 1 \end{pmatrix}.$$

For $1 \le i_1 < i_2 < i_3 < i_4 \le r$ and $1 \le j_1 < j_2 < j_3 \le r$ define the object shape coordinates

$$m_{i_1,i_2,i_3,i_4} = \det \begin{pmatrix} x_{i_1} & x_{i_2} & x_{i_3} & x_{i_4} \\ y_{i_1} & y_{i_2} & y_{i_3} & y_{i_4} \\ z_{i_1} & z_{i_2} & z_{i_3} & z_{i_4} \\ 1 & 1 & 1 & 1 \end{pmatrix}$$

and the image shape coordinates

$$n_{j_1,j_2,j_3} = \det \begin{pmatrix} u_{i_1} & u_{i_2} & u_{i_3} \\ v_{i_1} & v_{i_2} & v_{i_3} \\ 1 & 1 & 1 \end{pmatrix}$$

Then the points $Q_1, \ldots, Q_r$ are the images of $P_1, \ldots, P_r$ under a generalized weak perspective projection if and only if

$$\sum_{1 \le \lambda_1 < \lambda_2 \le r} \epsilon_{\lambda_1,\lambda_2} m_{\alpha_1,\alpha_2,\lambda_1,\lambda_2} n_{\gamma_1,\gamma_2,\gamma_3} = 0$$

for all choices of $1 \le \alpha_1 < \alpha_2 \le r$ and $1 \le \beta_1 < \beta_2 < \ldots < \beta_{r-5} \le r$ where $1 \le \gamma_1 < \gamma_2 < \gamma_3 \le r$ is the complement of $\{\lambda_1, \lambda_2, \beta_1, \ldots, \beta_{r-5}\}$ in $\{1, \ldots, r\}$ when $\lambda_1, \lambda_2, \beta_1, \ldots, \beta_{r-5}$ are distinct (otherwise $n_{\gamma_1,\gamma_2\gamma_3} = 0$) and $\epsilon_{\lambda_1,\lambda_2}$ is the sign of the permutation

$$\gamma_1, \gamma_2, \gamma_3, \lambda_1, \lambda_2, \beta_1, \ldots, \beta_{r-5}$$

of the numbers $1, \ldots, r$. The expressions $m_{\alpha_1,\alpha_2,\lambda_1,\lambda_2}$ and $n_{\gamma_1,\gamma_2,\gamma_3}$ should be treated as skew-symmetric in their indices.

As an example, consider the case $r = 5$. We pick $\alpha_1 = 1, \alpha_2 = 2$ and no β's are required. Our formula becomes

$$\sum_{1 \le \lambda_1 < \lambda_2 < 5} \epsilon_{\lambda_1,\lambda_2} m_{1,2,\lambda_1,\lambda_2} n_{\gamma_1\gamma_2\gamma_3}$$

when $\gamma_1\gamma_2\gamma_3$ is the complement of λ_1, λ_2 in $\{1, \ldots, 5\}$. This yields

$$m_{1234} n_{125} - m_{1235} n_{124} + m_{1245} n_{123} = 0.$$

We get 10 such equations as we vary α_1 and α_2.

2. THE FULL PERSPECTIVE CASE

We now turn our attention from generalized weak perspective projections to the so-called pinhole camera model, which is simply projection from a point P in projective space $\mathbb{P}^3$ onto a hyperplane H not containing P:

$$\pi : \mathbb{P}^3 - \{P\} \longrightarrow H \cong \mathbb{P}^2$$

This case becomes much more difficult since we are now considering configurations of points in projective space and hence are allowed to scale each of our points (homogeneous coordinates) by an arbitrary nonzero constant.

We will consider r points in projective 3-space, which we will again think of as feature points on an object. (There is a hyperplane that does not pass through any of these points and the complement of that hyperplane in $\mathbb{P}^3$ is isomorphic to $\mathbb{R}^3$.) We will refer to this set of points as a *projective object configuration* or simply an *object configuration* when it is clear that we are dealing with points in projective space. Now "taking a picture" of the object is just projecting the object configuration from a point onto a hyperplane (which is isomorphic to $\mathbb{P}^2$). We refer to this type of projection as a *full perspective projection*, and we call the image of a projective object configuration under such a projection a *projective image configuration* or simply an *image configuration*.

When we view projection from a point as a map $\pi : \mathbb{P}^3 \to \mathbb{P}^2$, our projections take the form

$$\begin{pmatrix} R \\ S \\ T \end{pmatrix} = T \begin{pmatrix} X \\ Y \\ Z \\ W \end{pmatrix}$$

where T is a 3×4 matrix of rank 3 and equality is in the sense of homogeneous coordinates. Conversely, every 3×4 matrix T of rank 3 defines a projection from some point in this way. More precisely, this point is given by the 1-dimensional null space of T (remember that points in projective n-space are 1-dimensional subspaces of affine $n+1$-space).

We should note that if $Q = (R : S : T) \in \mathbb{P}^2$ is the image of $P = (X : Y : Z : W) \in \mathbb{P}^3$ under a full perspective projection T (so $Q = TP$) then for any 3×3 scalar matrix A and any 4×4 scalar matrix B we have $Q = (ATB)P$. Thus, the set of full perspective projections is equivalent to the set of 3×4 matrices of rank 3 up to multiplication on the left or right by a scalar matrix.

Now let T be a full perspective projection. Let A be any 3×3 matrix with $\det(A) \neq 0$ and let B be any 4×4 matrix with $\det(B) \neq 0$. Then ATB is again a 3×4 matrix of rank 3, i.e. ATB is again a full perspective projection. Note that, as previously observed, if we multiply A and B by scalar matrices, the projection ATB remains unchanged as a map between projective spaces. Thus, we should view A as an element of $PGL(3)$ and B as an element of $PGL(4)$. (In general, $PGL(k)$ is the quotient $GL(k)/S$ where S is the subgroup of scalar matrices.)

The impact here is that the best we can hope to do is to relate object configurations up to a $PGL(4)$ transformation with image configurations up to a $PGL(3)$ transformation. Hence, our object shape space should be $U/PGL(4)$ for some open set $U \subset (\mathbb{P}^3)^r$ and our image shape space will be $W/PGL(3)$for some open set $W \subset (\mathbb{P}^2)^r$, when we have r point features.

2.1. The Associated Variety of a Configuration

In the affine case, we were able to assign to each shape a distinct point in a fixed projective space. Unfortunately in the full perspective case, our ability to scale the homogeneous coordinates of the points of our configurations complicates matters, so that no convenient analogue of the affine shape coordinates are available. We circumvent this problem by instead assigning to each configuration a natural projective variety. Later in this paper, we will discuss the possibilities made available by using Chow forms to give global coordinates on our projective shape spaces.

Although ultimately we want to consider configurations of r points in $\mathbb{P}^2$ and $\mathbb{P}^3$, let us begin by examining configurations of 4 points in $\mathbb{P}^1$. Let $P_i = (x_i : y_i) \in \mathbb{P}^1$ for $1 \leq i \leq 4$. We will assume that the points are not all the same point. In the spirit of the affine case, we make this configuration with these representative homogeneous coordinates into a matrix

$$\mathcal{M} = \begin{pmatrix} x_1 & x_2 & x_3 & x_4 \\ y_1 & y_2 & y_3 & y_4 \end{pmatrix}$$

As in the affine case, this matrix corresponds to the point

$$(m_{12} : m_{13} : m_{14} : m_{23} : m_{24} : m_{34}) \in G(2,4) \subset \mathbb{P}^{\binom{4}{2}-1} = \mathbb{P}^5 \text{ where } m_{ij} = \det \begin{pmatrix} x_i & x_j \\ y_i & y_j \end{pmatrix} \quad 1 \leq i < j \leq 4,$$

noting that since the points are not identical in $\mathbb{P}^1$, at least one of the m_{ij} is nonzero.

If for each $1 \leq i \leq 4$ we scale P_i by a nonzero constant a_i, we have the same configuration, but our matrix is now

$$\begin{pmatrix} x_1 & x_2 & x_3 & x_4 \\ y_1 & y_2 & y_3 & y_4 \end{pmatrix} \begin{pmatrix} a_1 & 0 & 0 & 0 \\ 0 & a_2 & 0 & 0 \\ 0 & 0 & a_3 & 0 \\ 0 & 0 & 0 & a_4 \end{pmatrix}$$

which corresponds to the point

$$(a_1a_2m_{12} : a_1a_3m_{13} : a_1a_4m_{14} : a_2a_3m_{23} : a_2a_4m_{24} : a_3a_4m_{34}) \in G(2,4) \subset \mathbb{P}^5$$

Thus for a given configuration of 4 points in $\mathbb{P}^1$ we have a map $\Phi : (\mathbb{R}^*)^4 \to G(2,4)$ given by

$$\Phi(a_1, a_2, a_3, a_4) = (a_1a_2m_{12} : a_1a_3m_{13} : a_1a_4m_{14} : a_2a_3m_{23} : a_2a_4m_{24} : a_3a_4m_{34})$$

(here $\mathbb{R}^*$ is the multiplicative group of nonzero elements of $\mathbb{R}$). Notice however that

$$\Phi(a,a,a,a) = a^2(m_{12} : m_{13} : m_{14} : m_{23} : m_{24} : m_{34}) = (m_{12} : m_{13} : m_{14} : m_{23} : m_{24} : m_{34}) \text{ in } \mathbb{P}^5.$$

So we have in fact a well defined map $\overline{\Phi} : (\mathbb{R}^*)^4/\mathbb{R}^* \cong (\mathbb{R}^*)^3 \to G(2,4)$ whose image we will denote $\mathcal{V}(P_1, P_2, P_3, P_4)$ $\subset G(2,4) \subset \mathbb{P}^5$ (or simply $\mathcal{V}$ when the configuration we are working with is understood). Thus, to each configuration we may assign a variety $\overline{\mathcal{V}(P_1, P_2, P_3, P_4)}$, the closure of $\mathcal{V}$ in $\mathbb{P}^5$, which we will call the *associated variety* of the configuration.

PROPOSITION 2.1. *Every configuration P_1, P_2, P_3, P_4 is assigned a unique variety $\mathcal{V}(P_1, P_2, P_3, P_4)$, and if two configurations P_1, P_2, P_3, P_4 and P_1', P_2', P_3', P_4' have the same associated variety, then they differ by a $PGL(2)$ transformation (and hence give the same point in our shape space).*

Proof. The fact that every configuration is assigned a unique variety is obvious. So suppose that for two configurations $P_i = (x_i : y_i)$, $1 \leq i \leq 4$ and $P_i' = (x_i' : y_i')$, $1 \leq i \leq 4$ we have $\mathcal{V}(P_1, P_2, P_3, P_4) = \mathcal{V}(P_1', P_2', P_3', P_4')$. Then for some $a_1, a_2, a_3, a_4 \in \mathbb{R}^*$

$$(m_{12} : m_{13} : m_{14} : m_{23} : m_{24} : m_{34}) = (a_1a_2m_{12}' : a_1a_3m_{13}' : a_1a_4m_{14}' : a_2a_3m_{23}' : a_2a_4m_{24}' : a_3a_4m_{34}')$$

where $m_{ij} = \det \begin{pmatrix} x_i & x_j \\ y_i & y_j \end{pmatrix}$ and $m_{ij}' = \det \begin{pmatrix} x_i' & x_j' \\ y_i' & y_j' \end{pmatrix}$. So we have that the null spaces of the matrices

$$\begin{pmatrix} x_1 & x_2 & x_3 & x_4 \\ y_1 & y_2 & y_3 & y_4 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} a_1x_1' & a_2x_2' & a_3x_3' & a_4x_4' \\ a_1y_1' & a_2y_2' & a_3y_3' & a_4y_4' \end{pmatrix}$$

give the same point under the Plücker embedding and hence are in fact the same linear subspace of $\mathbb{R}^4$. Thus the matrices differ by the left action of a $GL(2)$ matrix from which we see that the configurations P_1, P_2, P_3, P_4 and P_1', P_2', P_3', P_4' differ by a $PGL(2)$ transformation. □

Now, having placed our configurations P_1, P_2, P_3, P_4 (up to a $PGL(2)$ transformation) in one-to-one correspondence with the projective varieties $\overline{\mathcal{V}(P_1, P_2, P_3, P_4)}$, we would like to understand the relations that the points in $\mathcal{V}$ must satisfy. So let $P_1, P_2, P_3, P_4 \in \mathbb{P}^1$ and let $(x_{12} : x_{13} : x_{14} : x_{23} : x_{24} : x_{34})$ be a point in $\mathcal{V} = \mathcal{V}(P_1, P_2, P_3, P_4)$. Then for some $a_1, a_2, a_3, a_4 \in \mathbb{R}^*$ the following must hold

$$\begin{aligned} x_{12} - a_1a_2m_{12} &= 0 \\ x_{13} - a_1a_2m_{13} &= 0 \\ x_{14} - a_1a_4m_{14} &= 0 \\ x_{23} - a_2a_3m_{23} &= 0 \\ x_{24} - a_2a_4m_{24} &= 0 \\ x_{34} - a_3a_4m_{34} &= 0 \end{aligned}$$

Using Gröebner bases, we eliminate the a_i's from this system and obtain the following Theorem

THEOREM 2.2. *$\overline{\mathcal{V}}$ is the zero locus of three polynomials in the variables $x_{12}, \ldots, x_{34}$*

$$f_1 = m_{12}m_{34}x_{13}x_{24} - m_{13}m_{24}x_{12}x_{34}$$
$$f_2 = m_{12}m_{34}x_{14}x_{23} - m_{14}m_{23}x_{12}x_{34}$$
$$f_3 = m_{13}m_{24}x_{14}x_{23} - m_{14}m_{23}x_{13}x_{24}$$

These same relations can also be obtained by observing that if i_1, i_2, i_3, i_4 and j_1, j_2, j_3, j_4 are two permutations of 1,2,3,4 then

$$\frac{m_{i_1i_2}m_{i_3i_4}x_{j_1j_2}x_{j_3j_4}}{m_{j_1j_2}m_{j_3j_4}x_{i_1i_2}x_{i_3i_4}} = \frac{m_{i_1i_2}m_{i_3i_4}(a_{j_1}a_{j_2}m_{j_1j_2})(a_{j_3}a_{j_4}m_{j_3j_4})}{m_{j_1j_2}m_{j_3j_4}(a_{i_1}a_{i_2}m_{i_1i_2})(a_{i_3}a_{i_4}m_{i_3i_4})} = 1$$

We should note that since $(m_{12} : m_{13} : m_{14} : m_{23} : m_{24} : m_{34})$ and $(x_{12} : x_{13} : x_{14} : x_{23} : x_{24} : x_{34})$ are points in $G(2,4) \subset \mathbb{P}^5$, the Plücker relations

$$p_1 = m_{12}m_{34} - m_{13}m_{24} + m_{14}m_{23} = 0$$
$$p_2 = x_{12}x_{34} - x_{13}x_{24} + x_{14}x_{23} = 0$$

are satisfied. It is easily seen from these relations that as long as enough of the m_{ij} are nonzero, we have that $V(f_1) = V(f_2) = V(f_3)$ as subvarieties of $G(2,4)$ and hence $\overline{\mathcal{V}}$ is defined as the zero locus of any *one* of f_1, f_2, f_3. In particular $\overline{\mathcal{V}}$ is a hypersurface in $G(2,4)$ and so has dimension $\dim(\mathcal{V}) = \dim(G(2,4))$-1=3.

All of the preceding discussion can be easily generalized to the case of $r \geq n+2$ points in $\mathbb{P}^n$. For each configuration $P_i = (x_{i,0}, \ldots, x_{i,n})$, $1 \leq i \leq r$ of r points in $\mathbb{P}^n$, we have a map $\overline{\Phi} : (\mathbb{R}^*)^r/\mathbb{R}^* \to G(n+1, r)$ obtained by constructing a matrix whose columns are representatives of $P_1, \ldots, P_r$ in $\mathbb{P}^n$ and then scaling the columns of that matrix. We denote the image of $\overline{\Phi}$ by $\mathcal{V}(P_1, \ldots, P_r)$. Thus, we place the configurations $P_1, \ldots, P_r$ of r points in $\mathbb{P}^n$ in one-to-one correspondence with the projective varieties $\overline{\mathcal{V}(P_1, \ldots, P_r)}$.

Explicitly, the map $\overline{\Phi} : (\mathbb{R}^*)^r/\mathbb{R}^* \to G(n+1, r)$ is given by

$$\overline{\Phi}(a_1, \ldots, a_r) = (a_{I_1}m_{I_1} : \ldots : a_{I_N}m_{I_N})$$

where $I_1, \ldots, I_N$ are the $(n+1)$-subsets of $\{1, \ldots, r\}$ and $a_{I_k} = \prod_{j \in I_k} a_j$. We would like to know for which configurations is $\overline{\Phi}$ one-to-one. In other words, we would like to know for which configurations do we have

$$(a_{I_1}m_{I_1} : \ldots : a_{I_N}m_{I_N}) = (m_{I_1} : \ldots : m_{I_N}) \qquad \Leftrightarrow \qquad a_i = a_j \text{ for all } i, j$$

The following theorem gives a large set of configurations (but not necessarily all) for which $\overline{\Phi}$ is 1-1.

THEOREM 2.3. *Suppose $P_1, \ldots, P_r$ is a configuration of r points in $\mathbb{P}^n$ so that there is a subset $P_{i_1}, \ldots, P_{i_{n+2}}$ of $n+2$ points in this configuration having the following properties:*

1. *for every subset $J = \{j_1, \ldots, j_{n+1}\} \subset \{i_1, \ldots, i_{n+2}\}$ the points $P_{j_1}, \ldots, P_{j_{n+1}}$ do not lie in a single hyperplane (i.e. $m_J \neq 0$)*
2. *there is some subset $K = \{k_1, \ldots, k_n\} \subset \{i_1, \ldots, i_{n+2}\}$ such that for all P_s not in the set $\{P_{i_1}, \ldots, P_{i_{n+2}}\}$ we have that the points $P_{k_1}, \ldots, P_{k_n}, P_s$ do not all lie in a single hyperplane (i.e. $m_{k_1 \ldots k_n s} \neq 0$).*

Then, the map $\overline{\Phi}$ is injective.

Proof. We will show that under these conditions,

$$(a_{I_1}m_{I_1} : \ldots : a_{I_N}m_{I_N}) = (m_{I_1} : \ldots : m_{I_N}) \qquad \Leftrightarrow \qquad a_i = a_j \text{ for all } i, j.$$

Note that

$$(a_{I_1}m_{I_1} : \ldots : a_{I_N}m_{I_N}) = (m_{I_1} : \ldots : m_{I_N})$$

if and only for all $i \neq j$,

$$\frac{a_{I_i}m_{I_i}}{a_{I_j}m_{I_j}} = \frac{m_{I_i}}{m_{I_j}}$$

assuming of course that $m_{I_j} \neq 0$.

First, let $\alpha, \beta \in \{1, \ldots, r\}$ be such that α, β are not in the set $\{i_1, \ldots i_{n+2}\}$. Then by condition 2, if we let $A = k_1, \ldots, k_n, \alpha$ and let $B = k_1, \ldots, k_n, \beta$ we have that $m_A \neq 0$ and $m_B \neq 0$. Thus since

$$\frac{a_A m_A}{a_B m_B} = \frac{m_A}{m_B}$$

we have that

$$\frac{a_\alpha}{a_\beta} = 1.$$

and hence $a_\alpha = a_\beta$.

Now, let α, β be such that α is in $\{i_1, \ldots, i_{n+2}\}$ but β is not. Choose $j_1, \ldots, j_n$ in $\{i_1, \ldots, i_{n+2}\}$ so that $j_s \neq j_t$ if $s \neq t$ and so that $j_s \neq \alpha$ for all s. Let $A = \{j_1, \ldots, j_n, \alpha\}$ and let $B = \{k_1, \ldots, k_n, \beta\}$. Then by condition 1, $m_A \neq 0$ and by condition 2, $m_B \neq 0$. Thus as above, we again get that $a_\alpha = a_\beta$.

A similar argument shows that if α and β are both in $\{i_1, \ldots, i_{n+2}\}$ then $a_\alpha = a_\beta$. Thus, under conditions 1 and 2, we have that the map $\overline{\Phi}$ is 1-1. □

We see now that for configurations $P_1, \ldots, P_r$ satisfying conditions (1) and (2), $\mathcal{V}(P_1, \ldots, P_r)$ is isomorphic to $(\mathbb{R}^*)^r/\mathbb{R}^* \cong \mathbb{R}^{r-1}$. In particular, $\dim\left(\overline{\mathcal{V}(P_1, \ldots, P_r)}\right) = r - 1$ which is consistent with our result in the case of 4 points in $\mathbb{P}^1$.

We do have a slight variation from the case of 4 points in $\mathbb{P}^1$ when we compute the defining equations of $\mathcal{V}(P_1, \ldots, P_r)$. Consider the case of 5 points $P_1, \ldots, P_5$ in $\mathbb{P}^1$. Then we observe that for a point $(x_{12} : \ldots : x_{45})$ in $\mathcal{V}(P_1, \ldots, P_5)$ we have for some $a_1, \ldots, a_5 \in \mathbb{R}^*$

$$\frac{m_{12}m_{13}m_{45}x_{14}x_{13}x_{25}}{m_{14}m_{13}m_{25}x_{12}x_{13}x_{45}} = \frac{m_{12}m_{13}m_{45}(a_1a_4x_{14})(a_1a_3x_{13})(a_2a_5x_{25})}{m_{14}m_{13}m_{25}(a_1a_2x_{12})(a_1a_3x_{13})(a_4a_5x_{45})} = 1$$

giving us the relation

$$m_{12}m_{13}m_{45}x_{14}x_{13}x_{25} - m_{14}m_{13}m_{25}x_{12}x_{13}x_{45} = 0.$$

So in general we will have some repetition of the entries of the indices even though in the case of 4 points in $\mathbb{P}^1$ we did not.

THEOREM 2.4. *For a configuration $P_1, \ldots, P_r$ of r points in $\mathbb{P}^n$, the variety $\overline{\mathcal{V}(P_1, \ldots, P_r)}$ is the zero locus of the following system of polynomials*

$$m_{I_1}m_{I_2}\cdots m_{I_k}x_{J_1}x_{J_2}\cdots x_{J_k} - m_{J_1}m_{J_2}\cdots m_{J_k}x_{I_1}x_{I_2}\cdots x_{I_k}$$

where $I_1, \ldots, I_k, J_1, \ldots, J_k$ ranges over all $n+1$-subsets of $\{1, \ldots r\}$ with the property that $\bigcup_{i=1}^k I_i = \bigcup_{i=1}^k J_i$ as multisets and k ranges from 2 to some positive integer $N(r)$. The exact value of $N(r)$ is not known, but computation of some small examples seems to indicate that $N(r) = r - 2$.

3. THE PROJECTIVE OBJECT-IMAGE RELATIONS

Given a projective object configuration $P_1, \dots, P_r$ and a projective image configuration $Q_1, \dots, Q_r$, we want to (as in the affine case) find necessary and sufficient conditions for the Q_i to be a full perspective projection of the P_i. Since every object configuration (fixing its homogeneous coordinates) gives a point in $G(4, r) \subset \mathbb{P}^{\binom{r}{4}-1}$ and every image configuration (fixing its homogeneous coordinates) gives a point in $G(3, r) \subset \mathbb{P}^{\binom{r}{3}-1}$, the closure of the set of matching object-image pairs should be a projective variety defined by a system of bihomogeneous polynomials in the Plücker coordinates $m_{1234}, \dots, m_{r-3 \dots r}$ on $G(4, r)$ and the Plücker coordinates $n_{123}, \dots, n_{r-2 \dots r}$ on $G(3, r)$. These relations should be satisfied independent of our choice of representatives (homogeneous coordinates) for our object and image configurations. In other words, we should have that if an image configuration $Q_1, \dots, Q_r$ is a full perspective projection of an object configuration $P_1, \dots, P_r$ then the product variety $\mathcal{V}(P_1, \dots, P_r) \times \mathcal{V}(Q_1, \dots, Q_r)$ should be completely contained in V.

Now, consider an object configuration $P_1, \dots, P_r$ with P_1, P_2, P_3, P_4, P_5 in general position. We may then move the configuration by a projective transformation so that $P_1 = (1:0:0:0), P_2 = (0:1:0:0), P_3 = (0:0:1:0), P_4 = (0:0:0:1)$ and $P_5 = (1:1:1:1)$. Assume also that for all $i \geq 6$, P_i does not lie in the plane defined by P_1, P_2, P_3 so that $P_i = (p_{3i-17} : p_{3i-16} : p_{3i-15} : 1)$. It turns out that we can write $p_1, \dots, p_{3r-15}$ in terms of Plücker coordinates in the following way

$$p_{3i-17} = \frac{m_{234i} m_{1235}}{m_{123i} m_{2345}}, \quad p_{3i-16} = \frac{m_{134i} m_{1235}}{m_{123i} m_{1345}}, \quad p_{3i-15} = \frac{m_{124i} m_{1235}}{m_{123i} m_{1245}}.$$

Note that the p_j are defined independent of our choice of representatives of $P_1, \dots, P_r$ for if we scale each P_i by a nonzero constant a_i, we get

$$\begin{aligned}
p_{3i-17} &= \frac{(a_2 a_3 a_4 a_i m_{234i})(a_1 a_2 a_3 a_5 m_{1235})}{(a_1 a_2 a_3 a_i m_{123i})(a_2 a_3 a_4 a_5 m_{2345})} = \frac{m_{234i} m_{1235}}{m_{123i} m_{2345}} \\
p_{3i-16} &= \frac{(a_1 a_3 a_4 a_i m_{134i})(a_1 a_2 a_3 a_5 m_{1235})}{(a_1 a_2 a_3 a_i m_{123i})(a_1 a_3 a_4 a_5 m_{1345})} = \frac{m_{134i} m_{1235}}{m_{123i} m_{1345}} \\
p_{3i-15} &= \frac{(a_1 a_2 a_4 a_i m_{124i})(a_1 a_2 a_3 a_5 m_{1235})}{(a_1 a_2 a_3 a_i m_{123i})(a_1 a_2 a_4 a_5 m_{1245})} = \frac{m_{124i} m_{1235}}{m_{123i} m_{1245}}
\end{aligned}$$

The values $p_1, \dots, p_{3r-15}$ form a fundamental set of invariants for our object configuration.

Similarly let $Q_1, \dots, Q_r$ be an image configuration with Q_1, Q_2, Q_3, Q_4 in general position and such that for $i \geq 5$, Q_i is not on the line defined by Q_1 and Q_2. We move the configuration by a projective transformation so that $Q_1 = (1:0:0), Q_2 = (0:1:0), Q_3 = (0:0:1), Q_4 = (1:1:1)$ and for each $i \geq 5$, $Q_i = (q_{2i-9} : q_{2i-8} : 1)$. The projective invariants $q_1, \dots, q_{2n-8}$ are again defined independent of our choice of representatives and are given in Plücker coordinates as

$$q_{2i-9} = \frac{n_{23i} n_{124}}{n_{12i} n_{234}}, \quad q_{2i-8} = \frac{n_{13i} n_{124}}{n_{12i} n_{134}}.$$

When we make the preceding assumptions about the positioning of our configurations, the object-image equations have been completely determined [12]. For example, in the case where $n = 6$, we have only one object-image relation given in terms of the projective invariants:

$$-q_2 q_3 p_2 p_3 + q_3 p_2 p_3 - q_3 p_3 - q_1 q_4 p_1 - q_1 p_1 p_2 + q_1 p_1 = -q_1 q_4 p_1 p_3 + q_4 p_1 p_3 - q_4 p_3 - q_2 q_3 p_2 - q_2 p_1 p_2 + q_2 p_2.$$

Making the appropriate substitutions and then clearing denominators and removing monomial factors, we have the object-image relation in terms of the Plücker coordinates to be

$$\begin{aligned}
&n_{125}n_{136}n_{234}m_{1236}m_{1246}m_{1345}m_{2345} - n_{123}n_{136}n_{234}m_{1236}m_{1246}m_{1345}m_{2345}\\
&-n_{126}n_{135}n_{234}m_{1236}m_{1245}m_{1346}m_{2345} + n_{124}n_{135}n_{236}m_{1236}m_{1245}m_{1346}m_{2345}\\
&+n_{125}n_{134}n_{236}m_{1235}m_{1245}m_{1346}m_{2345} - n_{124}n_{135}n_{236}m_{1235}m_{1246}m_{1346}m_{2345}\\
&+n_{126}n_{134}n_{235}m_{1236}m_{1245}m_{1345}m_{2346} - n_{124}n_{136}n_{235}m_{1236}m_{1245}m_{1345}m_{2346}\\
&-n_{125}n_{136}n_{234}m_{1235}m_{1246}m_{1345}m_{2346} + n_{124}n_{136}n_{235}m_{1235}m_{1246}m_{1345}m_{2346}\\
&+n_{126}n_{135}n_{234}m_{1235}m_{1245}m_{1346}m_{2346} - n_{126}n_{134}n_{235}m_{1235}m_{1245}m_{1346}m_{2346} = 0
\end{aligned}$$

We should note that since the p_i and q_i are defined independent of our choice of representatives for the P_i and Q_i, this relation will be satisfied independent of our choice of representatives.

Now let σ be a permutation of $1,\ldots,r$. Suppose that in our object configuration $P_1,\ldots,P_r$ the points $P_{\sigma(1)}, P_{\sigma(2)}, P_{\sigma(3)}, P_{\sigma(4)}, P_{\sigma(5)}$ are in general position and that for all $k \geq 6$, $P_{\sigma(k)}$ is not in the span of $P_{\sigma(1)}, P_{\sigma(2)}, P_{\sigma(3)}$. Then we may move our configuration by a projective transformation so that $P_{\sigma(1)} = (1:0:0:0), P_{\sigma(2)} = (0:1:0:0), P_{\sigma(3)} = (0:0:1:0), P_{\sigma(4)} = (0:0:0:1), P_{\sigma(5)} = (1:1:1:1)$, and for $k \geq 6$, $P_{\sigma(k)} = (p'_{3k-17} : p'_{3k-16} : p'_{3k-15} : 1)$.

Similarly, let τ be a permutation of $1,\ldots,n$, and suppose that in our image configuration $Q_1,\ldots,Q_r$ the points $Q_{\tau(1)}, Q_{\tau(2)}, Q_{\tau(3)}, Q_{\tau(4)}$ are in general position and that for all $k \geq 6$, $Q_{\tau(k)}$ is not in the span of $Q_{\tau(1)}$and $Q_{\tau(2)}$. We now move $Q_1,\ldots,Q_r$ by a projective transformation so that $Q_{\tau(1)} = (1:0:0), Q_{\tau(2)} = (0:1:0), Q_{\tau(3)} = (0:0:1), Q_{\tau(4)} = (1:1:1)$ and for $k \geq 5$, $Q_{\tau(k)} = (q'_{2k-9} : q'_{2k-8} : 1)$.

We now have a new set of object invariants $p'_1,\ldots,p'_{3r-15}$ and a new set of image invariants $q'_1,\ldots,q'_{2r-8}$ which, as before, may be written in terms of Plücker coordinates

$$p'_{3i-17} = \frac{m_{\sigma(2)\sigma(3)\sigma(4)\sigma(i)}m_{\sigma(1)\sigma(2)\sigma(3)\sigma(5)}}{m_{\sigma(1)\sigma(2)\sigma(3)\sigma(i)}m_{\sigma(2)\sigma(3)\sigma(4)\sigma(5)}}$$

$$p'_{3i-16} = \frac{m_{\sigma(1)\sigma(3)\sigma(4)\sigma(i)}m_{\sigma(1)\sigma(2)\sigma(3)\sigma(5)}}{m_{\sigma(1)\sigma(2)\sigma(3)\sigma(i)}m_{\sigma(1)\sigma(3)\sigma(4)\sigma(5)}}$$

$$p'_{3i-15} = \frac{m_{\sigma(1)\sigma(2)\sigma(4)\sigma(i)}m_{\sigma(1)\sigma(2)\sigma(3)\sigma(5)}}{m_{\sigma(1)\sigma(2)\sigma(3)\sigma(i)}m_{\sigma(1)\sigma(2)\sigma(4)\sigma(5)}}$$

$$q'_{2i-9} = \frac{n_{\tau(2)\tau(3)\tau(i)}n_{\tau(1)\tau(2)\tau(4)}}{n_{\tau(1)\tau(2)\tau(i)}n_{\tau(2)\tau(3)\tau(4)}}$$

$$q'_{2i-8} = \frac{n_{\tau(1)\tau(3)\tau(i)}n_{\tau(1)\tau(2)\tau(4)}}{n_{\tau(1)\tau(2)\tau(i)}n_{\tau(1)\tau(3)\tau(4)}}$$

keeping in mind that we view the m_{ijkl} and the n_{stu} as skew-symmetric in their indices.

Using the method of [12] we get a new set of object-image relations in terms of the new invariants which we may again write in terms of Plücker coordinates. We should notice that since our projective transformations are completely determined by sending $P_{\sigma(1)}, P_{\sigma(2)}, P_{\sigma(3)}, P_{\sigma(4)}, P_{\sigma(5)}$ to $(1:0:0:0), (0:1:0:0), (0:0:1:0), (0:0:0:1), (1:1:1:1)$ respectively and by sending $Q_{\tau(1)}, Q_{\tau(2)}, Q_{\tau(3)}, Q_{\tau(4)}$ to $(1:0:0), (0:1:0), (0:0:1), (1:1:1)$ respectively, we may assume that $\sigma(6) < \ldots < \sigma(r)$ and that $\tau(5) < \ldots < \tau(r)$. Taking all of these object-image relations as σ ranges over all permutations of $1,\ldots,r$ with $\sigma(6) < \ldots < \sigma(r)$ and as τ ranges over all permutations of $1,\ldots,r$ with $\tau(5) < \ldots < \tau(r)$ gives us a global system of object-image relations. This system is still grossly overdetermined and more work is being done to reduce the number of relations in this system.

4. CONCLUSION

The next step is to give a concrete description of the shape spaces in the full perspective case. This will mean collapsing the associated variety $\overline{\mathcal{V}(P_1, \ldots, P_r)}$ to a point. One way to do this is via the Chow form and the Chow point of $\overline{\mathcal{V}}$ (see [8]). We would then realize the shape space as a quasi-projective variety in some projective space where it will acquire a natural metric. This program is the subject of our current work. While the object-image equations provide a test for matching, the metrics provide an even more robust approach to matching. For example, we often want to know if two configurations of a fixed number of points in 2D or 3D are the same if we allow projective transformations. If they are, then we want a distance of zero, and if not, we want a distance that expresses their dissimilarity – always recognizing that we can transform the points. The Procrustes metric, described in the shape theory literature [6] and [7], provides such a notion of distance for similarity transformations. However, it does not work for perspective transformations. Moreover, it is fixed in a particular dimension. By that we mean that it cannot be regarded as giving us a notion of "distance" between, say, a 3D configuration of points and a 2D configuration of points, where zero distance corresponds to the 2D points being a full perspective projection of the 3D points. However, the metrics we developed in the affine case can be used to give a natural measure of object-image matching. These metrics also provide a rigorous foundation for error and statistical analysis in the object recognition problem. Similar metrics can be derived in the full perspective case using the approach mentioned. The details will be in our forthcoming papers.

REFERENCES

1. G. Arnold, P.F. Stiller, and K. Sturtz, "Object-image metrics for generalized weak perspective projection," chapter in *Statistics and Analysis of Shape*, Hamid Krim (ed.), Birkhäuser, pp. 253–279, 2006.
2. G. Arnold, P.F. Stiller, and K. Sturtz, "Geometric methods for ATR – invariants, object-image equations, and metrics," July 2003, preprint.
3. Harris, *Algebraic Geometry*, Graduate Text in Mathematics, **133**, Springer-Verlag, 1992.
4. W.V.D. Hodge and D. Pedoe, *Methods of Algebraic Geometry*, nos. 1, 2, and 3, in *Mathematical Library Series*, Cambridge University Press, 1994.
5. B. Howard, J. Millson, A. Snowden, and R. Vakil, "The projective invariants of ordered points on the line," preprint, Feb. 2006.
6. D.G. Kendall, "Shape manifolds, Procrustean metrics, and complex projective space," *Bull. London Math. Soc.* **16**, pp. 81–121, 1984.
7. K.V. Mardia and I.L. Dryden, *Statistical Shape Analysis*, John Wiley and Sons, Chichester, England, 1998.
8. I.R. Shafarevich, *Basic Algebraic Geometry*, Die Grundlehren der mathematischen Wissenschaften, Band 213, Springer-Verlag, New York, 1974.
9. P.F. Stiller, "The relationship between shape under similarity transformations and shape under affine transformations," in *Proc. SPIE Int'l Symposium on Optical Science and Technology, Mathematics of Data/Image Coding, Compression and Encryption, with Applications*, Mark Schmalz (ed.), Vol. 5561, pp. 108–116, Denver, CO, Aug. 2004.
10. P.F. Stiller, "Vision metrics and object/image relations II: Discrimination metrics and object/image duality," in *Proc. SPIE Int'l Conf., Electronic Imaging, Vision Geometry XII*, Vol. 5300, pp 74–85, San Jose, CA, Jan. 2004.
11. P.F. Stiller, "Object/image relations, shape spaces, and metrics: The generalized weak perspective case," preprint, 2004.
12. P.F. Stiller and B. Huber, "Geometric hashing and object recognition," in *Proc. SPIE Int'l Conf., Vision Geometry VIII*, Vol. 3811, pp. 201–211, Denver, CO, 1999.

Mathematical Aspects of Shape Analysis for Object Recognition

D. Gregory Arnold[a)] and Peter F. Stiller[b)]
a) Air Force Research Lab, AFRL/SNAT, Bldg. 620
2241 Avionics Circle, Dayton, Ohio 45433–7302
b) Department of Mathematics, MS3368, Texas A&M University
College Station, Texas 77843-3368

Abstract

In this paper we survey some of the mathematical techniques that have led to useful new results in shape analysis and their application to a variety of object recognition tasks. In particular, we will show how these techniques allow one to solve a number of fundamental problems related to object recognition for configurations of point features under a generalized weak perspective model of image formation. Our approach makes use of progress in shape theory and includes the development of object-image equations for shape matching and the exploitation of shape space metrices (especially object-image metrics) to measure matching up to certain transformations. This theory is built on advanced mathematical techniques from algebraic and differential geometry which are used to construct generalized shape spaces for various projection and sensor models. That construction in turn is used to find natural metrics that express the distance (geometric difference) between two configurations of object features, two configurations of image features, or an object and an image pair. Such metrics are believed to produce the most robust tests for object identification; at least as far as the object's geometry is concerned. Moreover, these metrics provide a basis for efficient hashing schemes to do identification quickly, and they provide a rigorous foundation for error and statistical analysis in any recognition system. The most important feature of a shape theoretic approach is that all of the matching tests and metrics are independent of the choice of coordinates used to express the feature locations on the object or in the image. In addition, the approach is independent of the camera/sensor position and any camera/sensor parameters. Finally, the method is also independent of object pose or image orientation. This is what makes the results so powerful.

Keywords: shape analysis, object recognition, shape space, generalized weak perspective, affine group, shape coordinates, object-image metric, Riemannian metric.

1 Introduction

A solution to the problem of single-view recognition is often a crucial first step in many target recognition and computer vision tasks. Understanding how information available in a single image of an object, be it an optical image, a SAR image, or a radar range profile, relates to the target object's geometry is a key step in building reliable identification algorithms. For example, without a priori knowledge of a sensor's viewpoint, an object's pose, or a sensor's parameters, it is difficult to efficiently recognize a three-dimensional arrangement of features (such as a geometric configuration of lines and/or points) on an object from a single two dimensional view. What is needed is an approach that is invariant to changing viewpoints, adjustments in the sensor parameters, or changes in the object's pose. Unfortunately, existing methods all too often rely on computationally expensive template matching that is, strictly speaking, neither view nor pose invariant. Specifically, those methods use comparisons against templates created for each possible view and pose;

with the infinite range of possibilities being approximated by some finite number of discrete views. Fortunately, recent mathematical developments in the theory of shape provide an alternative. To carry out such an invariant, shape theoretic approach to target recognition, we need to seek out and exploit properties and relationships that are geometrically intrinsic to the objects and/or images being compared. Moreover, to develop this approach for different types of sensors, we must take into account the fact that each type requires a different model of image formation and therefore a different form of invariance. Radar and Ladar sensors require the use of an orthographic or scaled orthographic model, while most optical sensors will use either a weak perspective, a generalized weak perspective, or a full perspective model.

Once we understand, for various sensors, the contribution of a single image toward the recognition or recovery of the geometry/shape of the object, it becomes easier to develop methods to integrate the information from multiple images taken by uncalibrated, distributed sensors of varying types, or to make use of a series of images taken by a single sensor of a moving object. It also makes it easier to understand and create flexible algorithms adapted to situations where the objects are not rigid but more deformable, as is the case with many of the recognition problems related to biometric or medical applications (e.g. face recognition, detecting heart or tissue anomalies, gait recognition, etc.)

The requirement of view and pose invariance, as well as the desirability of a coordinate independent formulation, leads us to start with a characterization of a configuration of object or image features by its 3D, 2D, or 1D shape, a mathematical notion related to geometric invariance. The specific transformation group (Euclidean group, similarity or conformal group, affine group, or projective general linear group) to which things should be invariant will be a function of the sensor type. We then need a fundamental set of equations that expresses the relationship between the 3D geometry (shape) and its "residual" in a 2D (or 1D) image. These are known as object-image equations. They completely and invariantly describe the mutual 3D/2D (or 1D) constraints. These equations can be exploited in a number of ways. For example, from a given 2D configuration, one can determine a set of non-linear constraints on the shape (geometric invariants) of the 3D configurations capable of producing that given 2D configuration, and thus arrive at a test for determining the object being viewed. Conversely, given a 3D geometric configuration (features on an object), one can dervive a set of equations that constrain the shape of the images of that object; helping to determine if that particular object appears in selected images.

The ultimate goal in all cases is to improve on and develop new algorithms for target recognition. Our approach in this paper uses advanced mathematical techniques from algebraic and differential geometry to construct generalized shape spaces for various projection and sensor models and formulates the object-image equations in terms of the global shape coordinates for these spaces. We then use the natural metrics on the shape spaces (which provide a measure of dissimilarity between two object configurations or two image configurations up to the allowed transformations) to find natural object-image metrics that express the distance (failure to match) between an object-image pair. Zero value for these metrics will mean matching up to the relevant transformations and/or projections. These metrics are pose and view invariant and are expressed in coordinate free terms. They produce the most robust tests for target identification; at least as far as target geometry is concerned. Moreover, such metrics provide the basis for efficient hashing schemes to do target identification quickly and also provide a rigorous foundation for error and statistical analysis in the ATR process.

Because of the limited space we have, we will content ourselves with introducing these ideas in the generalized weak perspective case, which models an optical sensor where the object is in the far field of view. This case is the most mathematically tractable and complete. Details can be found in the references. For now, we introduce the theory and give some examples.

2 The Generalized Weak Perspective (Affine) Case

We consider r points in space, which we think of as feature points on some object. We refer to this set of points as an *object configuration*. Next, we "take a picture" of the object by choosing a plane and projecting these feature points into that plane. We will call this set of points in the plane an *image configuration*.

In this section we will (1) identify the space of shapes, which are configurations modulo the action of a certain group of transformations on $\mathbb{R}^n$, $n = 2, 3$, and give global coordinates on the shape space, (2) give necessary and sufficient conditions for an image configuration to be a projection of an object configuration, and (3) define a natural metric on the shape spaces. For additional details see Arnold, Stiller, and Sturtz [2].

2.1 The Generalized Weak Perspective Projection

The type of projections we will consider are called *generalized weak perspective projections*. If we represent points P in $\mathbb{R}^n$ ($n = 2$ or 3) in column form $P = (x_1, \ldots, x_n, 1)^T$, then these projections take the form of a linear map T from $\mathbb{R}^3$ to $\mathbb{R}^2$ given by a matrix

$$T = \begin{pmatrix} t_{11} & t_{12} & t_{13} & t_{14} \\ t_{21} & t_{22} & t_{23} & t_{14} \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

where T has rank 3.

Now let A be an invertible 3×3 matrix and let B be an invertible 4×4 matrix. It turns out that if T is a generalized weak perspective projection, then ATB is a generalized weak perspective projection if an only if A and B are *affine transformations* i.e. A and B take the form

$$\begin{pmatrix} & & & c_1 \\ & S & & \vdots \\ & & & c_n \\ 0 & \cdots & 0 & 1 \end{pmatrix}$$

where $S \in GL(n)$ and $c_1, \ldots, c_n \in \mathbb{R}$.

What does this mean in terms of our object and image configurations? Suppose $Q \in \mathbb{R}^2$ is the image of a point $P \in \mathbb{R}^3$ under a generalized weak perspective projection T, i.e. $Q = TP$. Then if we move P by some affine transformation B to another point P' and if we move Q to another point Q' by an affine transformation A, we will have that $Q' = ATB^{-1}P'$. As a result, we see that Q' is the image of P' under the generalized weak perspective projection ATB^{-1} (since A and B^{-1} are both affine transformations).

This observation shows us that by choosing to consider generalized weak perspective projections, the best that we can hope to do is relate object configurations to image configurations up to affine transformations.

2.2 The Affine Shape Spaces

As the preceding observation suggests, we should consider two configurations (object or image) equivalent if they differ by an affine transformation. In a sense equivalent configurations are the same object or image just rotated, translated, scaled, or otherwise moved by an affine transformation. Alternatively, we can view equivalent configurations as being the same object or image, but with their feature locations expressed in a different coordinate system. We would like to construct the space of configurations of r points in $\mathbb{R}^n$ modulo the action of the group of affine transformations. These spaces would then represent the distinct objects and images independent of pose or view. To do this we must assume that the points in our configuration are non-coplanar for $n = 3$ or non-collinear for $n = 2$, which is reasonable since a configuration of coplanar points in $\mathbb{R}^3$ would in fact be a configuration of points in $\mathbb{R}^2$ and would not represent a real 3D object, etc.

Let $P_i = (x_{i,1}, \ldots, x_{i,n})$ for $i = 1 \ldots r$, $r \geq n+2$ be a configuration of r non-coplanar (or non-collinear) points in $\mathbb{R}^n$, $n = 3$ (or 2), and consider the matrix

$$M = \begin{pmatrix} x_{1,1} & x_{2,1} & & x_{r,1} \\ x_{1,2} & x_{2,2} & \cdots & x_{r,2} \\ \vdots & \vdots & & \vdots \\ x_{1,n} & x_{n,2} & & x_{r,n} \\ 1 & 1 & & 1 \end{pmatrix}$$

Now to the configuration $P_1, \ldots, P_r$ we associate an $(r-n-1)$-dimensional linear subspace, $K^{r-n-1} \subset \mathbb{R}^r$. In particular, K^{r-n-1} is the null space of M when we view M as a linear map from $\mathbb{R}^r$ to $\mathbb{R}^{n+1}$. The fact that K^{r-n-1} has dimension $r-n-1$ follows from the observation that M has rank $n+1$ as a linear map because at least one $(n+1) \times (n+1)$ minor of M has non-zero determinant due to non-coplanarity (or non-collinearity).

The important thing to notice is that if we apply an affine transformation A to our configuration we obtain a new $(n+1) \times r$ matrix $M' = AM$, but the null space of M' is exactly K^{r-n-1}, the null space of M. Moreover, since $K^{r-n-1} \subset H^{r-1} = \{(v_1, \ldots, v_r) \in \mathbb{R}^r \mid \sum_{i=1}^r v_i = 0\}$, we may assign to our configuration the unique point $[K^{r-n-1}] \in G(r-n-1, H^{r-1})$, the Grassmannian of $(r-n-1)$-dimensional subspaces in the $(r-1)$-dimensional space H^{r-1}, a well understood compact manifold of dimension $n(r-n-1)$.

Definition 2.1. We call the manifold $X = G(r-n-1, H^{r-1})$ the *affine shape space* for configurations of r points in $\mathbb{R}^n$. If $n = 3$, we will call $X = G(r-4, r-1)$ *affine object space* (or just *object space*) and refer to points in this space as *object shapes.* If $n = 2$, we will call $X = G(r-3, r-1)$ *affine image space* (or just *image space*) and refer to its elements as *image shapes.*

Every point in X is of the form $[K^{r-n-1}]$ for some configuration $P_1, \ldots, P_r \in \mathbb{R}^n$, and most importantly, if two configurations $P_1, \ldots, P_r \in \mathbb{R}^n$ and $P'_1, \ldots, P'_r \in \mathbb{R}^n$ give the same point in X, then they differ by an affine transformation.

2.3 The Plücker Embedding

Since X is a real manifold, we can find local coordinates for a point $[K] \in X$; however, since we ultimately want to give relations that tell us when an image configuration is a projection of an object configuration, it would be more convenient to find global coordinates on X. We may do so by mapping X into a projective space via the *Plücker embedding.* In general, the Plücker embedding embeds a Grassmannian $G(n, r)$ (n-dimensional subspaces of an r-dimensional vector space, V^r) in the projective space $\mathbb{P}\left(\bigwedge^{r-n} V^r\right) \cong \mathbb{P}^{\binom{r}{r-n}-1} \cong \mathbb{P}^{\binom{r}{n}-1}$ as a projective variety in the following way: let $[K] \in G(n, r)$. Then K is the intersection of $r-n$ hyperplanes in our vector space V^r, where each hyperplane is given by a linear form

$$0 = \sum_{i=1}^{r} k_{j,i} \widehat{e_i}, \quad j = 1, \ldots, r-n$$

where $e_1, \ldots, e_r$ is a basis for V^r and $\widehat{e_i}$ are the dual basis. More simply put, K is the null space of the matrix

$$L = \begin{pmatrix} k_{1,1} & k_{1,2} & \ldots & k_{1,r} \\ k_{2,1} & k_{2,2} & \ldots & k_{2,r} \\ \vdots & \vdots & \ddots & \vdots \\ k_{r-n,1} & k_{r-n,2} & \ldots & k_{r-n,r} \end{pmatrix}$$

Now for each $1 \leq i_1 < i_2 < \ldots < i_{r-n} \leq r$ we define $[i_1, i_2, \ldots, i_{r-n}]$ to be the determinant of the $(r-n) \times (r-n)$ minor of L whose columns are the $i_1, i_2, \ldots, i_{r-n}$ columns of L, i.e.

$$[i_1, i_2, \ldots, i_{r-n}] = \det \begin{pmatrix} k_{1,i_1} & k_{1,i_2} & \ldots & k_{1,i_{r-n}} \\ k_{2,i_1} & k_{2,i_2} & \ldots & k_{2,i_{r-n}} \\ \vdots & \vdots & \ddots & \vdots \\ k_{r-n,i_1} & k_{r-n,i_2} & \ldots & k_{r-n,i_{r-n}} \end{pmatrix}$$

The Plücker embedding is now defined to be the map

$$\begin{array}{rccll} \Phi_{n,r}: & G(n,r) & \longrightarrow & \mathbb{P}^{\binom{r}{n}-1} & \\ & [K] & \longmapsto & ([1,2,\ldots,r-n] : \ldots : [n+1,n+2,\ldots,r]) & \text{(all minors)} \end{array}$$

and the homogeneous coordinates of $\Phi_{n,r}([K])$ are called the *Plücker coordinates* of K.

It is important to note that this map does not depend on our choice of hyperplanes, but does depend on our choice of basis for V^r. We should also note that this map does in fact embed $G(n,r)$ as a closed projective variety in $\mathbb{P}^{\binom{r}{n}-1}$. In other words, $\Phi_{n,r}(G(n,r))$ is the zero locus of some system of homogeneous polynomials $f_1, \ldots, f_s$ in the variables $x_{1,2,\ldots,r-n}; \ldots; x_{n+1,\ldots,r}$ with coefficients in the base field of V^r. We use the variables $x_{1,2,\ldots,r-n}; \ldots; x_{n+1,\ldots,r}$ to indicate that the $x_{i_1,i_2,\ldots,i_{r-n}}$ coordinate of $\Phi_{n,r}([K])$ is $[i_1, \ldots, i_{r-n}]$. The equations $f_i = 0,\ 1 \leq i \leq s$ are known as the Plücker relations (see [4] or [5]).

One way to give global coordinates on $X = G(r-n-1, H^{r-1})$, would be to embed X into the projective space $\mathbb{P}_{\mathbb{R}}^{\binom{r-1}{n}-1}$ via the Plücker embedding $\Phi_{r-n-1,r-1}$. However, this would require us to choose a basis for H^{r-1}. Fortunately, there is a very natural way to avoid this problem.

Since $K^{k-n-1} \subset H^{r-1} \subset \mathbb{R}^r$ we may view X as a submanifold of $G(r-n-1,r)$, in which case $\Phi_{r-n-1,r}$ embeds X in $\mathbb{P}_{\mathbb{R}}^{\binom{r}{n+1}-1}$ as a subvariety of $\Phi_{r-n-1,r}(G(r-n-1,r))$. Under this map, a configuration $P_i = (x_{i,1}, \ldots, x_{i,n}),\ i = 1, \ldots, r$ is mapped into $\mathbb{P}_{\mathbb{R}}^{\binom{r}{n+1}-1}$ by taking all the determinants of the maximal minors of our original feature point matrix M.

Embedding our shape space X into $\mathbb{P}_{\mathbb{R}}^{\binom{r}{n+1}-1}$ in this fashion is in some sense a more natural way to give global coordinates on X than embedding it into $\mathbb{P}_{\mathbb{R}}^{\binom{r-1}{n}-1}$. This method allows us to work directly with the matrix determined by our configuration rather than forcing us to choose a basis for H^{r-1} and then rewriting our basis for K^{r-n-1} in terms of our chosen basis for H.

Definition 2.2. Given a configuration $P_1, \ldots, P_r \in \mathbb{R}^n$ we will refer to the Plücker coordinates of K^{r-n-1} viewed as a subspace of $\mathbb{R}^r$ (rather than H^{r-1}) as the *shape coordinates* of the configuration $P_1, \ldots, P_r$.

2.4 The Object/Image Relations

Given an object configuration $P_1, \ldots, P_r$ and an image configuration $Q_1, \ldots, Q_r$ we want to give necessary and sufficient conditions (the object-image relations) for the Q_i to be a generalized weak perspective projection of the P_i. Recall that we view our object space X as a subvariety of $\mathbb{P}^{\binom{r}{4}-1}$ and our image space Y as a subvariety of $\mathbb{P}^{\binom{r}{3}-1}$. As such, we want to view the set V of pairs (K, L) where L is an image shape that comes from a generalized weak perspective projection of the object shape K (the so-called set of matching object-image pairs) as a subvariety $V \subset X \times Y \subset \mathbb{P}^{\binom{r}{4}-1} \times \mathbb{P}^{\binom{r}{3}-1}$. Therefore, our object-image relations should be a system of bihomogeneous polynomials in the object and image shape coordinates whose zero locus is precisely V.

Recall that our object shapes are linear subspaces $K^{r-4} \subset \mathbb{R}^r$ of dimension $r-4$ and our image shapes are linear subspaces $L^{r-3} \subset \mathbb{R}^n$ of dimension $r-3$. The following relates object and image shapes under generalized weak perspective projection.

Theorem 2.3. *Let $P_1, \ldots, P_r$ be an object configuration with corresponding object shape K^{r-4} and let $Q_1, \ldots, Q_r$ be an image configuration with corresponding image shape L^{r-3}. Then the Q_i are a generalized weak perspective projection of the P_i if and only if*

$$K^{r-4} \subset L^{r-3} \subset H^{r-1} \subset \mathbb{R}^r$$

This fact and the incidence relations given in Theorem I, §5, Chapter VII of Hodge and Pedoe [5] give us our object-image relations.

Theorem 2.4. *Let $P_i = (x_i, y_i, z_i)$, $1 \leq i \leq r$ be an object configuration with corresponding matrix*

$$M = \begin{pmatrix} x_1 & x_2 & & x_r \\ y_1 & y_2 & \cdots & y_r \\ z_1 & z_2 & & z_r \\ 1 & 1 & & 1 \end{pmatrix}$$

and let $Q_i = (u_i, v_i)$, $1 \leq i \leq r$ be an image configuration with corresponding matrix

$$N = \begin{pmatrix} u_1 & u_2 & & u_r \\ v_1 & v_2 & \cdots & v_r \\ 1 & 1 & & 1 \end{pmatrix}.$$

For $1 \leq i_1 < i_2 < i_3 < i_4 \leq r$ and $1 \leq j_1 < j_2 < j_3 \leq r$ define the object shape coordinates

$$m_{i_1,i_2,i_3,i_4} = \det \begin{pmatrix} x_{i_1} & x_{i_2} & x_{i_3} & x_{i_4} \\ y_{i_1} & y_{i_2} & y_{i_3} & y_{i_4} \\ z_{i_1} & z_{i_2} & z_{i_3} & z_{i_4} \\ 1 & 1 & 1 & 1 \end{pmatrix}$$

and the image shape coordinates

$$n_{j_1,j_2,j_3} = \det \begin{pmatrix} u_{i_1} & u_{i_2} & u_{i_3} \\ v_{i_1} & v_{i_2} & v_{i_3} \\ 1 & 1 & 1 \end{pmatrix}$$

Then the points $Q_1, \ldots, Q_r$ are the images of $P_1, \ldots, P_r$ under a generalized weak perspective projection if and only if

$$\sum_{1 \leq \lambda_1 < \lambda_2 \leq r} \epsilon_{\lambda_1,\lambda_2} m_{\alpha_1,\alpha_2,\lambda_1,\lambda_2} n_{\gamma_1,\gamma_2,\gamma_3} = 0$$

for all choices of $1 \leq \alpha_1 < \alpha_2 \leq r$ *and* $1 \leq \beta_1 < \beta_2 < \ldots < \beta_{r-5} \leq r$ *where* $1 \leq \gamma_1 < \gamma_2 < \gamma_3 \leq r$ *is the complement of* $\{\lambda_1, \lambda_2, \beta_1, \ldots, \beta_{r-5}\}$ *in* $\{1, \ldots, r\}$ *when* $\lambda_1, \lambda_2, \beta_1, \ldots, \beta_{r-5}$ *are distinct (otherwise* $n_{\gamma_1,\gamma_2\gamma_3} = 0$*) and* $\epsilon_{\lambda_1,\lambda_2}$ *is the sign of the permutation* $\gamma_1, \gamma_2, \gamma_3, \lambda_1, \lambda_2, \beta_1, \ldots, \beta_{r-5}$ *of the numbers* $1, \ldots, r$*. The expressions* $m_{\alpha_1,\alpha_2,\lambda_1,\lambda_2}$ *and* $n_{\gamma_1,\gamma_2,\gamma_3}$ *should be treated as skew-symmetric in their indices.*

As an example, consider the case $r = 5$. We pick $\alpha_1 = 1, \alpha_2 = 2$ and no β's are required. Our formula becomes

$$\sum_{1 \leq \lambda_1 < \lambda_2 < 5} \epsilon_{\lambda_1,\lambda_2} m_{1,2,\lambda_1,\lambda_2} n_{\gamma_1\gamma_2\gamma_3}$$

where $\gamma_1\gamma_2\gamma_3$ is the complement of λ_1, λ_2 in $\{1, \ldots, 5\}$. This yields $m_{1234}n_{125} - m_{1235}n_{124} + m_{1245}n_{123} = 0$. We get 10 such equations as we vary α_1 and α_2:

$$\begin{array}{ll}
0 = m_{1234}n_{124} - m_{1235}n_{124} + m_{1245}n_{123} & 0 = m_{1234}n_{245} - m_{1245}n_{234} + m_{2345}n_{124} \\
0 = m_{1234}n_{135} - m_{1235}n_{134} + m_{1345}n_{123} & 0 = m_{1235}n_{245} - m_{1245}n_{235} + m_{2345}n_{125} \\
0 = m_{1234}n_{145} - m_{1245}n_{134} + m_{1345}n_{124} & 0 = m_{1234}n_{345} - m_{1345}n_{234} + m_{2345}n_{134} \\
0 = m_{1235}n_{145} - m_{1245}n_{135} + m_{1345}n_{125} & 0 = m_{1235}n_{345} - m_{1345}n_{235} + m_{2345}n_{135} \\
0 = m_{1234}n_{235} - m_{1235}n_{234} + m_{2345}n_{123} & 0 = m_{1245}n_{345} - m_{1345}n_{245} + m_{2345}n_{145}
\end{array}$$

3 Metrics

How far apart are two object shapes or two image shapes? Since the shape spaces are Grassmannians, we can use the natural Riemannian metric on these manifolds, known as the Fubini–Study metric to define distances (see Arias, Edelman, and Smith [1]).

3.1 A Riemannian Metric on the Object Shape Space and on the Image Shape Space

Given two objects, i.e. two r-tuples $P_1, \ldots, P_r$ and $\widetilde{P}_1, \ldots, \widetilde{P}_r$ of points in $\mathbb{R}^3$. We define the distance between them, or more specifically, the distance between their shapes K^{r-4} and $\widetilde{K}^{r-4}$, as follows. First we choose orthonormal bases for K^{r-4} and $\widetilde{K}^{r-4}$ as subspaces of $\mathbb{R}^r$ and arrange those vectors as the columns of two $r \times (r-4)$ orthonormal matrices K and $\widetilde{K}$. We then compute the singular values of the $(r-4) \times (r-4)$ matrix $\widetilde{K}^T K$ and denote by θ_i, $(i = 1, \ldots, r-4)$ the arc cosines of the singular values. These angles are the so-called principal angles between the subspaces.

Definition 3.1. The affine shape distance in object space between two r-tuples of object feature points is defined to be

$$d_{\mathrm{Obj}}(K^{r-4}, \widetilde{K}^{r-4}) = \sqrt{\sum_{i=1}^{r-4} \theta_i^2} = \sqrt{\sum_{i=1}^{r-4} (\arccos \lambda_i)^2}$$

where λ_i are the singular values of $\widetilde{K}^T K$ for the orthonormal matrices K and $\widetilde{K}$ created by choosing orthonormal bases of the subspaces K^{r-4} and $\widetilde{K}^{r-4}$ in $\mathbb{R}^r$. See below for examples.

Definition 3.2. Given two r-tuples of points $Q_1, \ldots, Q_r$ and $\widetilde{Q}_1, \ldots, \widetilde{Q}_r$ in the plane representing certain image features, we define the affine shape distance in image space between them to be

$$d_{\mathrm{Im}}(L^{r-3}, \widetilde{L}^{r-3}) = \sqrt{\sum_{j=1}^{r-3} \varphi_j^2} = \sqrt{\sum_{j=1}^{r-3} (\arccos \tau_i)^2}$$

where τ_i are the singular values of $\widetilde{L}^T L$ for orthonormal matrices L and $\widetilde{L}$ created by choosing orthonormal bases for L^{r-3} and $\widetilde{L}^{r-3}$ in $\mathbb{R}^r$. See below for examples.

We remark that these distances are the natural metric distances on the shape spaces X and Y which are the Grassmann manifolds $Gr_{\mathbb{R}}(r-4, H^{r-1})$ and $Gr_{\mathbb{R}}(r-3, H^{r-1})$, because they are geodesic submanifolds of $Gr_{\mathbb{R}}(r-4, r)$ and $Gr_{\mathbb{R}}(r-3, r)$ respectively.

3.2 A "Metric" Measure of Matching Between an Object Feature Set and an Image Feature Set

Finally, we can compute a "distance" between an object $[K^{r-4}] \in X$ and an image $[L^{r-3}] \in Y$. This can be done in two ways. First working in object space X, we get

$$d^1_{O/I}([K^{r-4}], [L^{r-3}]) = \min_{\widetilde{K}^{r-4}} d_{\mathrm{Obj}}(K^{r-4}, \widetilde{K}^{r-4})$$

where $\widetilde{K}^{r-4}$ runs over all objects capable of producing image L^{r-3}, i.e. all subspaces $\widetilde{K}^{r-4} \subset L^{r-3}$. Second working in image space Y, we get

$$d^2_{O/I}([K^{r-4}], [L^{r-3}]) = \min_{\widetilde{L}^{r-3}} d_{\mathrm{Im}}(L^{r-3}, \widetilde{L}^{r-3})$$

where $\widetilde{L}^{r-3}$ runs over all images of the object K^{r-4}, i.e. all subspaces $\widetilde{L}^{r-3} \subset H^{r-1}$ which contain K^{r-4}.

In both cases these values work out to be the square root of the sum of the squares of the principal angles between K^{r-4} and L^{r-3} computed from the arc-cosines of the singular values of $L^T K$ in the same manner as above.

Theorem 3.3 (Object/Image Metric Duality). *The distance between a set of object features $P_1 \ldots P_r$ and a set of image features $Q_1 \ldots Q_r$ can be computed either in object space by minimizing the affine shape distance between $P_1, \ldots, P_r$ and all object r-tuples which are capable of being projected to $Q_1, \ldots, Q_r$ (via a generalized weak perspective projection), or in image space as the minimum affine shape distance between $Q_1, \ldots, Q_r$ and all generalized weak perspective projections of $P_1, \ldots, P_r$. Moreover, these two minimums are equal, i.e. $d^1_{O/I} = d^2_{O/I}$.*

Of course $d_{O/I} = 0$ if and only if $P_1, \ldots, P_r$ can be projected to $Q_1, \ldots, Q_r$ via a generalized weak perspective projection. We remark that analogs of this result can be proved for other projection models.

4 Examples

In this section we create a number of examples and provide the Mathematica code necessary to implement some of the results of the paper. Additional code to generate the shape space equations and the object-image equations can be obtained from the author. The examples below involve the case of five ($n = 5$) feature points projected from 3D to 2D under generalized weak perspective projection (also known as the affine case).

4.1 Object Data

We begin by creating three (arbitrary) 3D objects. The objects are described by point features written as columns of a 4 by n matrix in so-called homogeneous form $(x, y, z, 1)$. Also the determinants of its 4 by 4 minors (the so-called Plücker or shape coordinates) are listed lexicographically. Some of the syntax here is taken from Mathematica commands.

$$\text{ObjectDataW} = \begin{pmatrix} 2 & 1 & -1 & 3 & -2 \\ 1 & -\frac{1}{2} & \frac{1}{2} & 0 & \frac{3}{2} \\ 1 & \frac{3}{2} & 2 & \frac{3}{2} & \frac{7}{4} \\ 1 & 1 & 1 & 1 & 1 \end{pmatrix} \qquad \text{Minors[ObjectDataW, 4]} = \{\{\tfrac{11}{4}, -\tfrac{7}{4}, -\tfrac{27}{8}, -\tfrac{7}{8}, -2\}\}$$

$$\text{ObjectDataX} = \begin{pmatrix} 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & -1 \\ 0 & 0 & 1 & 1 & \frac{1}{2} \\ 1 & 1 & 1 & 1 & 1 \end{pmatrix} \qquad \text{Minors[ObjectDataX, 4]} = \{\{-2, \tfrac{3}{2}, \tfrac{5}{2}, \tfrac{1}{2}, \tfrac{3}{2}\}\}$$

$$\text{ObjectDataY} = \begin{pmatrix} 1 & \frac{3}{2} & 2 & 0 & -3 \\ -1 & \frac{1}{2} & 0 & 0 & -1 \\ 0 & -\frac{1}{2} & 1 & 0 & 2 \\ 1 & 1 & 1 & 1 & 1 \end{pmatrix} \qquad \text{Minors[ObjectDataY, 4]} = \{\{3, 10, -2, -8, 1\}\}$$

4.2 Group Actions

We now use rotation, translation, scale and reflection matrices to create object data equivalent to (i.e. having the same shape as) ObjectDataY above.

$$\text{RotAndTrans1} = \begin{pmatrix} \frac{1}{3} & -\frac{2}{3} & \frac{2}{3} & 1 \\ \frac{2}{3} & -\frac{1}{3} & -\frac{2}{3} & -1 \\ \frac{2}{3} & \frac{2}{3} & \frac{1}{3} & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \qquad \text{ScaleAndReflec1} = \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$\text{ObjectDataZ1} = \text{RotAndTrans1.ObjectDataY} = \begin{pmatrix} 2 & \frac{5}{6} & \frac{7}{3} & 1 & 2 \\ 0 & \frac{1}{6} & -\frac{1}{3} & -1 & -4 \\ 0 & \frac{7}{6} & \frac{5}{3} & 0 & -2 \\ 1 & 1 & 1 & 1 & 1 \end{pmatrix}$$

$$\text{Minors[ObjectDataZ1, 4]} = \{\{3, 10, -2, -8, 1\}\}$$

$$\text{ObjectDataZ2} = \text{ScaleAndReflec1.RotAndTrans1.ObjectDataY} = \begin{pmatrix} 4 & \frac{5}{3} & \frac{14}{3} & 2 & 4 \\ 0 & \frac{1}{2} & -1 & -3 & -12 \\ 0 & -\frac{7}{6} & -\frac{5}{3} & 0 & 2 \\ 1 & 1 & 1 & 1 & 1 \end{pmatrix}$$

$$\text{Minors[ObjectDataZ2, 4]} = \{\{-18, -60, 12, 48, -6\}\}$$

Note that ObjectDataY, ObjectDataZ1, and ObjectDataZ2 all have, up to scale, the same 5-vector of Plücker (shape) coordinates, because they are just transformations of the same object.

4.3 Projections

Now we create some generalized weak perspective projection matrices taking us from 3D to 2D.

$$\text{projection1} = \begin{pmatrix} 1 & -2 & 0 & 1 \\ -1 & 0 & 1 & 3 \\ 0 & 0 & 0 & 1 \end{pmatrix} \qquad \text{projection2} = \begin{pmatrix} \frac{1}{2} & -1 & -1 & 2 \\ 2 & \frac{3}{2} & -2 & \frac{1}{2} \\ 0 & 0 & 0 & 1 \end{pmatrix} \qquad \text{projection3} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Finally, we form a number of images using these projections and the object data above.

$$\text{ImageData1X} = \text{projection1.ObjectDataX} = \begin{pmatrix} 2 & -1 & 1 & 0 & 3 \\ 2 & 3 & 4 & 3 & \frac{7}{2} \\ 1 & 1 & 1 & 1 & 1 \end{pmatrix}$$

$$\text{Minors[ImageData1X, 3]} = \begin{pmatrix} -5 & -1 & -\frac{11}{2} & 3 & -\frac{7}{2} & -4 & -1 & -3 & \frac{1}{2} & \frac{5}{2} \end{pmatrix}$$

$$\text{ImageData2X} = \text{projection2.ObjectDataX} = \begin{pmatrix} \frac{5}{2} & 1 & 1 & \frac{1}{2} & \frac{5}{2} \\ \frac{5}{2} & 2 & -\frac{3}{2} & 2 & -2 \\ 1 & 1 & 1 & 1 & 1 \end{pmatrix}$$

$$\mathrm{Minors}[\mathrm{ImageData2X}, 3] = \begin{pmatrix} \frac{21}{4} & -\frac{1}{4} & \frac{27}{4} & -\frac{29}{4} & \frac{27}{4} & 9 & -\frac{7}{4} & \frac{21}{4} & 2 & -5 \end{pmatrix}$$

$$\mathrm{ImageData1Y} = \mathrm{projection1.ObjectDataY} = \begin{pmatrix} 4 & \frac{3}{2} & 3 & 1 & 0 \\ 2 & 1 & 2 & 3 & 8 \\ 1 & 1 & 1 & 1 & 1 \end{pmatrix}$$

$$\mathrm{Minors}[\mathrm{ImageData1Y}, 3] = \begin{pmatrix} -1 & -\frac{11}{2} & -19 & -1 & -6 & -14 & \frac{7}{2} & 12 & -\frac{1}{2} & -9 \end{pmatrix}$$

$$\mathrm{ImageData2Y} = \mathrm{projection2.ObjectDataY} = \begin{pmatrix} \frac{7}{2} & \frac{11}{4} & 2 & 2 & -\frac{1}{2} \\ 1 & \frac{21}{4} & \frac{5}{2} & \frac{1}{2} & -11 \\ 1 & 1 & 1 & 1 & 1 \end{pmatrix}$$

$$\mathrm{Minors}[\mathrm{ImageData2Y}, 3] = \begin{pmatrix} \frac{21}{4} & \frac{27}{4} & 26 & 3 & 24 & 16 & \frac{3}{2} & \frac{13}{4} & -\frac{13}{4} & -5 \end{pmatrix}$$

$$\mathrm{ImageData3W} = \mathrm{projection3.ObjectDataW} = \begin{pmatrix} 1 & -\frac{1}{2} & \frac{1}{2} & 0 & \frac{3}{2} \\ 1 & \frac{3}{2} & 2 & \frac{3}{2} & \frac{7}{4} \\ 1 & 1 & 1 & 1 & 1 \end{pmatrix}$$

$$\mathrm{Minors}[\mathrm{ImageData3W}, 3] = \begin{pmatrix} -\frac{5}{4} & -\frac{1}{4} & -\frac{11}{8} & \frac{3}{4} & -\frac{7}{8} & -1 & -\frac{1}{4} & -\frac{3}{4} & \frac{1}{8} & \frac{5}{8} \end{pmatrix}$$

Notice that ImageData1X and ImageData3W have the same image shape. They differ by a scale factor of 2 and the Plücker (shape) coordinates are the same up to a factor of 4. Thus we have an instance of two different, i.e. inequivalent, object shapes producing the same image shape. Likewise one can see that the images produced by a particular object, say ObjectDataX, can result in different (inequivalent) image shapes. For example ImageData1X and ImageData2X are not the same shape because their vector of shape coordinates are not scalar multiples of each other. This illustrates the many-to-one and one-to-many nature of the relationship between object shapes and image shapes.

4.4 Metrics

The next commands in Mathematica, while complicated looking, just create an orthonormal basis for the row span of our data matrices, multiplies two such matrices together, and then finds the singular values of the resulting 4 by 4, 3 by 3 or 3 by 4 (4 by 3) matrix. The arc-cosines of these singular values are the so-called principal angles between the subspaces spanned by the two sets of rows. The square root of the sum of the principle angles squared is the value of the metric. In the case of object space and of image space this is the distance provided by the natural Riemannian metric. Of course you are free to scale the metric. One obvious scaling is to set the total volume of the shape space equal to one, so that volume can be associated with a probability measure.

4.5 Object-Object Metric

What follows is a Mathematica function that computes the distance between two objects in object space. Note that the code doesn't depend on the number of feature points or the dimensionality of the points.

```
ObjectSpaceMetric[Obj1_, Obj2_]:=
 Norm[ArcCos[SingularValueList[N[QRDecomposition[Transpose[Obj1]][[1]].Transpose
   [QRDecomposition[Transpose[Obj2]][[1]]]], Tolerance → 0]]]
```

An an example lets compute some distances.

$$\mathrm{ObjectSpaceMetric}[\mathrm{ObjectDataW}, \mathrm{ObjectDataX}] = 0.0654569$$
$$\mathrm{ObjectSpaceMetric}[\mathrm{ObjectDataX}, \mathrm{ObjectDataY}] = 1.54176$$
$$\mathrm{ObjectSpaceMetric}[\mathrm{ObjectDataY}, \mathrm{ObjectDataZ2}] = 2.58096 \times 10^{-8}$$

Notice that this says ObjectDataW and ObjectDataX are fairly close but that ObjectDataX and ObjectDataY are relatively far apart. Of course ObjectDataY and ObjectDataZ2 are really zero distance apart because they differ by an affine transformation of 3-space.

4.6 Image-Image Metric

We now introduce the metric in image space. Note that it is given by the same code – only the input sizes have changed.

$$\text{ImageSpaceMetric[Im1\_, Im2\_]:=ObjectSpaceMetric[Im1, Im2];}$$

Let's now compute some distances in image space between our image shapes.

$$\begin{aligned}
&\text{ImageSpaceMetric[ImageData1X, ImageData1X]} = 0.\\
&\text{ImageSpaceMetric[ImageData1X, ImageData2X]} = 0.408449\\
&\text{ImageSpaceMetric[ImageData2X, ImageData1X]} = 0.408449\\
&\text{ImageSpaceMetric[ImageData2X, ImageData2X]} = 0.
\end{aligned}$$

The above computation illustrates that our metric is symmetric and of course gives zero distance between an image and itself. It also shows again that these two images of ObjectDataX are not equivalent and cannot be transformed one to the other by an affine transformation of 2-space.

$$\begin{aligned}
&\text{ImageSpaceMetric[ImageData1Y, ImageData2Y]} = 0.813806\\
&\text{ImageSpaceMetric[ImageData1X, ImageData1Y]} = 1.14686\\
&\text{ImageSpaceMetric[ImageData1X, ImageData2Y]} = 0.641144\\
&\text{ImageSpaceMetric[ImageData2X, ImageData1Y]} = 1.22873\\
&\text{ImageSpaceMetric[ImageData2X, ImageData2Y]} = 0.775138\\
&\text{ImageSpaceMetric[ImageData1X, ImageData3W]} = 0.
\end{aligned}$$

None of the images compared are equivalent, i.e. they are all distinct shapes, except for ImageData1X and ImageData3W which we previously observed were the same shape.

4.7 Object-Image Metric

Finally we introduce the object-image metric as a fundamental way to compare the matching of object data with image data.

$$\text{ObjectImageMetric[Obj\_, Im\_]} := \text{ObjectSpaceMetric[Obj, Im];}$$

This is in fact the metric discussed in the text, although a proof of that fact requires some work.

Examples:

$$\begin{aligned}
&\text{ObjectImageMetric[ObjectDataX, ImageData1Y]} = 1.14218\\
&\text{ObjectImageMetric[ObjectDataX, Imagedata1X]} = .49012 \times 10^{-8}\\
&\text{ObjectImageMetric[ObjectDataX, Imagedata2X]} = 2.10734 \times 10^{-8}
\end{aligned}$$

Note these later two are zero because the image data really is a generalized weak perspective projection of the object data The object-image metric will evaluate to zero if and only if a generalized weak perspective transformation exists which carries the object data to the image data.

Final example:

$$\text{ObjectImageMetric[ObjectDataX, ImageData3W]} = 1.49012 \times 10^{-8}$$
$$\text{ObjectImageMetric[ObjectDataW, ImageData1X]} = 2.10734 \times 10^{-8}$$
$$\text{ObjectImageMetric[ObjectDataW, ImageData2X]} = 0.0259829$$

Of course the first two of these are zero because imageData3W and imageData1X are both the same image shape and are projections of ObjectDataW and ObjectDataX respectively. However there is no way to project ObjectDataW to another of ObjectDataX's images (namely imageData2X) because our object-image "distance" in that case is not zero. **The closer the Object-Image metric is to zero the closer some projection of the given object will be to the given image.**

Object-Image Relations

Finally let's check an object-image equation. For ObjectDataX we have [1234] = -2 [1235] = 3/2 [1245] = 5/2. For ImageData1X we have [123] = -5 [124] = -1 [125] =-11/2. The first object-image equation is [1234][125]-[1235][124]+[1245][123]. The reader can check this is indeed zero. We leave it as an exercise to check the vanishing for the other nine object-image equations. (See above.) Note that knowing one image of an object imposes linear relations on the shape coordinates of the object and if we have sufficiently many independent views of the object, we can solve for its shape, which determines it uniquely up to an affine transformation of 3-space. Mathematica code is available to generate the object-image equations and the defining equations of the shape spaces inside the appropriate projective space.

References

[1] T. Arias, A. Edelman, and S. Smith, "The geometry of algorithms with orthogonality constraints," SIAM Journal of Matrix Analysis and Applications 20, 1998, pp. 303–353.

[2] G. Arnold, P.F. Stiller, and K. Sturtz, "Object-image metrics for generalized weak perspective projection," chapter in *Statistics and Analysis of Shape*, Hamid Krim (ed.), Birkhäuser, pp. 253–279, 2006.

[3] P. Griffiths and J. Harris, "Principles of Algebraic Geometry," Wiley and Sons, 1978.

[4] Harris, *Algebraic Geometry*, Graduate Text in Mathematics, **133**, Springer-Verlag, 1992.

[5] W.V.D. Hodge and D. Pedoe, "Methods of Algebraic Geometry," Volumes 1, 2, and 3, Cambridge Mathematical Library Series, Cambridge University Press, 1994.

[6] P.F. Stiller, "Vision metrics and object/image relations II: Discrimination metrics and object/image duality," Electronic Imaging, Vision Geometry XII, Vol. 5300, San Jose, CA, 1/04, pp. 74–85 (2004).

[7] P.F. Stiller, "Object/image relations and vision metrics I," Proceedings SPIE Int'l Symposium on Optical Science and Technology, Mathematics of Data/Image Coding, Compression and Encryption VI, with Applications, Vol. 5208, Mark Schmalz, Ed., San Diego, CA, 8/03, pp. 165–178 (2003).

[8] P.F. Stiller, "Object recognition from a global geometric perspective: invariants and metrics," Proceedings SPIE Int'l Conf., Vision Geometry XI, Vol. 4794, Seattle, WA, 7/02, pp. 71–80 (2002).